

## Redes e Serviços - Frame

### Introdução às Redes e Serviços

- Débito de transmissão = largura de banda
- Protocolos controlam o envio e receção de mensagens:
  - ↳ TCP, IP, HTTP, FTP, PPP
- Protocolo - definem o formato e a ordem das mensagens enviadas e recebidas entre as entidades da rede e as ações executadas quando de transmissão e receção de mensagens.

### Periferia da rede

- Estações (host) - contêm os programas com as aplicações. Exemplo: web, mail
- Modelo Cliente/Servidor - estação cliente solicita e recebe serviço de um servidor que está sempre à escuta. Exemplo: browser / servidor
- Modelo peer-to-peer - utilitário mínimo ou nula de servidores dedicados. Exemplo: skype, ...

### Core da Rede

- Os recursos são divididos em pedaços, estes são atribuídos às chamadas. Se o pedaço não for usado por nenhuma chamada, fica inativo
- Divisão da largura de banda em pedaços:
  - ↳ Divisão em frequência FDM
  - ↳ Divisão em tempo TDM
- Os pacotes são manipulados:
  - rotograma - os pacotes são tratados de forma independente, podem levar qualquer rota, chegam juntos de ordem
  - Circuitos Virtuais - estabelecidos entre pré-planeados, visados pacotes para o estabelecimento da ligação, cada pacote contém um identificador do circuito em vez do endereço destino.
- Redes de Acesso wireless - acesso wireless pointilhado liga estações ao router através de um ponto de acesso

### Redes Residenciais: componentes típicos

- ADSL ou modem de cabo
- Router / Firewall / NAT
- Ethernet
- Ponto de acesso wireless

## Modelo OSI

Nível 7 - Application - Aplicações / Serviços

Nível 6 - Presentation - Definição, manipulação e codificação de informações

Nível 5 - Session - Estabelecimento e manutenção de sessões

Nível 4 - Transporte - Comunicação extremo a extremo

Nível 3 - Network - Endereçamento e encaminhamento

Nível 2 - Data Link - Parilha do meio

Nível 1 - Physical - Transmissão dos Sinais

Pacotes - os pacotes incluem vários cabeçalhos concatenados

## LAN e Redes IP

- Redes locais (LAN) - interligam estações relativamente próximas através de ligações sanitárias
  - ↳ tecnologias - Ethernet, Token Ring, 802.11, FDDI, ...

Trama Ethernet - o adaptador de rede do emissor encapsula o datagrama IP no trama ethernet

Preamble End.Dest End.Origem Type Dados CRC

- Preamble - 7 bytes com o padrão 10101010 seguido de um byte 10101011 usado para sincronizar os reléios
- Endereços - 6 bytes - se o adaptador recebe um trama com o end. destino igual ao seu passa os dados da trama ao protocolo da camada de rede.
- Type - indica qual o protocolo da camada de rede. (tipicamente IP)
- CRC - detecção de erros de transmissão - se for detectado um o trama é descartada

### Endereços MAC (Endereço Físico) -

- função - permitir a transmissão de tramas entre interfaces ligadas fisicamente
- tem 48 bits - vem na RAM das placas de rede
- Notação Hexadecimal
- Cada placa de rede tem um endereço MAC (Broadcast: FF-FF-FF-FF-FF-FF)

### Endereçamento IP

- Endereço IP - identificado único de 4 bytes (32 bits)

- Interface - ligação entre host/roteador e a rede

Cada interface tem de ter pelo menos um endereço IP associado

## Classes de Endereços

Classe	Range Endereço	Range Endereço
A	1.0.0.0	128.0.0.0
B	128.0.0.0	191.255.0.0
C	192.0.0.0	223.255.255.0
D	224.0.0.0	239.255.255.255
E	240.0.0.0	255.255.255.255

## Endereços IP Especiais

- Tudo 0's - apenas válido durante inicializações
- Tudo 0's HostID - .....
- Tudo 1's - Broadcast local - não pode ser usado como endereço de origem
- NetID Tudo 1's - Broadcast dirigido à rede NetID
- 127 Qualquer - Loopback
- NetID Tudo 0's - Identificador da rede NetID

Máscaras - é utilizada para separar a parte de rede e a parte de host dos endereços

fronteiras flexíveis sendo estes definidos a partir de uma máscara

a parte da máscara (binária) cujo bit seja 1 → faz parte do identificador de rede  
os bits da máscara que sejam 0 → faz parte do identificador do host

CTDR - sub-divisão de redes em subconjuntos de endereços de comprimento arbitrário

Formato: a.b.c.d.(X) → nº de bits do identificador da (sub)-rede

/24 - 255.255.255.0

/25 - 255.255.255.128

/26 - 255.255.255.192

## Campos do datagrama IP

- Version -(4 bytes) versão do protocolo IP → versão actualmente mais comum 4
- Header Length -(4 bits) - tamanho do cabeçalho em blocos de 4 octetos
  - ↳ quando não tem opções cuja 5 bloco de octetos é o 1º octeto assume o valor 0x45
- Type of Service -(1 byte) - tipo de serviço ao qual o pacote pertence - padronizado 0x00
- Tamanho Total (2bytes) - tamanho do datagrama IP em octetos incluindo o cabeçalho
- Identification (2bytes) - identificador atribuído pela estação que gerou o datagrama
- Flags (3 bits)
  - 1 bit está reservado (assume sempre 0)

- o segundo bit assume o valor 0 se o datagrama puder ser desfragmentado e 1 caso contrário
- o terceiro bit assume o valor 0 se o último fragmento e 1 se
- Fragment offset (13 bits) - posição do fragmento no datagrama original (múltiplos de 8 bytes)
- Time to live (1 byte) - tempo máximo que o datagrama pode permanecer na rede
  - é alterado em cada roteador → quando chega o valor 0 é eliminado
- Protocol (1 byte) - especifica o protocolo de nível superior
- Header checksum (2 bytes) - resultado da soma (polinomial 16 bits) dos campos da cabecera

### ARP - Address Resolution Protocol

- É um protocolo usado para encontrar um endereço de comutador de ethernet, p.e., a partir de um endereço IP. O emissor difunde em broadcast um pacote ARP em que este contém o endereço IP de outro host e espera uma resposta com o endereço MAC respetivo. Cada máquina mantém uma tabela de resolução (ARP).
  - Quando uma máquina precisa de comunicar com outra por IP
  - precisa do seu endereço MAC, mas só sabe o IP
  - se o endereço da outra máquina não estiver na sua tabela ARP
  - a máquina que quer comunicar envia um pacote ARP Request em broadcast com o IP da máquina
  - todos os máquinas recebem
  - a máquina destino verifica que tem o endereço na sua tabela ARP e responde com ARP Response
  - imediatamente para a máquina que enviou o pedido
  - a máquina que precisa de comunicar vai actualizar a sua tabela ARP

- Esta resolução apenas existe dentro da mesma rede local.
- Pacotes ARP não "alcançam" roteadores

### Encaminhamento

Quando uma estação pretende enviar um pacote IP para uma rede IP que não a sua o primeiro salto é o default Gateway.

↳ configurado pelo utilizador e corresponde ao endereço IP da interface de um dos roteadores que pertence à rede da estação.

- Para enviar pacotes IP de uma máquina para outra roteador pede o procedimento é semelhante ao ARP mas quando a máquina que envia verifica que o endereço destino pertence a outra rede, envia o pacote via o seu gateway (roteador)
- Define o endereço MAC do seu gateway
- Constrói frame ethernet
- Encapsula o datagrama IP no frame ethernet e envia

## Tabelas de Encaminhamento

Destino (Identificador da rede)

Hop count

Next hop (próximo roteador de caminho)

Interface

Custo

Cada roteador tem uma tabela de encaminhamento com os endereços na forma descrita ali em cima. Descreve o caminho percorrido por uma mensagem desde o ponto origem até o seu ponto destino.

## Fragmentação e Reagrupamento em IP

Em cada rede existe um tamanho máximo de pacotes que podem ser transmitidos  $\rightarrow$  MTU

No caso de ser maior, o pacote é fragmentado à entrada e reagrupado à saída

Campos que intervêm na fragmentação: destination, fragment offset, flags

## IPv4 Subnetting

Endereço IPv4 - endereço IP único para um interface de rede

Exceções: endereços dinâmicos IPv4 (DHCP)

endereços IP em redes Privadas (NAT)

### ↳ Endereço IPv4

- identificador de 32 bits
- codifica um número de rede (prefixo de rede) e número de host

### Inteiro Decimal

Cada byte é identificado por um byte decimal no intervalo [0...255]

- Classe A - prefixo de rede 8 bits  $\rightarrow$  endereço começa por 0
- Classe B - prefixo de rede 16 bits  $\rightarrow$  endereço começa por 10
- Classe C - prefixo de rede 24 bits  $\rightarrow$  endereço começa por 110

- Os roteadores têm uma entrada na tabela de encaminhamento para cada endereço da rede
- Com o subnetting um roteador só precisa de saber uma entrada para cada classe
- Em 1993 o tamanho das tabelas de encaminhamento começou a superar a capacidade dos roteadores
- Consequência: a atribuição de classes teve de ser abandonada
- Benefícios:

- Para interpretação do espaço de endereços IP
- Restringir a atribuição de endereços IP para aumentar a eficiência
- Permite a agregação das rotas para minimizar entradas na tabela de encaminhamento

### Blocos de endereços

# endereços utilizáveis = nº de endereços de host - 2 endereços

Identificação da rede + endereço de Broadcast

### Atribuição de endereços com subnetting

A cada parte da rede é atribuído um intervalo de endereços IP

Os endereços em cada subrede podem ser administrados localmente

#### Vantagens:

- Hierarquia de 3 camadas: rede, sub-rede, host
- Reduz a complexidade do roteador, a complexidade das tabelas de encaminhamento do exterior é reduzida
- O comprimento da máscara de rede não precisa ser idêntico em todas as sub-redes.

### DHCP e DNS

#### DHCP - Dynamic Host Configuration Protocol (DHCP)

- Atribuição dinâmica de endereços IP a terminais
- Filosofia Cliente / Servidor
- Aluguel de endereços
- Configuração dos terminais com:
  - ↳ informações da máscara de rede
  - ↳ default gateway
  - ↳ servidores de DNS
  - ↳ servidores de WINS e domínio DNS

#### Configuração de um servidor DHCP

- Gama de Endereços - conjunto de endereços - endereço inicial e endereço final
- Gama de exclusão - conjunto de endereços que se querem excluir
- Endereços reservados - endereços IP atribuídos de forma permanente a endereços MAC
- Duração dos alugueres

## Protocolo DHCP

Extensão BOOTP - permite que um terminal sem disco descobre o seu endereço IP, um endereço de um servidor e o nome de um ficheiro a pedir ao servidor para ser copiado para a memória e executado localmente.

Protocolo de Aluguer → 4° fases:

### ↳ 1º fase - Discover

A mensagem DHCP Discover é encapsulada num pacote BootP request. Serve para descobrir os servidores de DHCP existentes na rede. O cliente pode indicar qual o endereço IP que pretende alugar.

### ↳ 2º fase - Offer

A mensagem DHCP Offer é encapsulada num pacote BootP Reply. Cada servidor indica um endereço IP para ser alugado (se possível os servidores respeitam a preferência do cliente).

### ↳ 3º fase - Request

A mensagem DHCP Request é encapsulada num pacote BootP Request. Após a negociação entre as possíveis diferentes ofertas recebidas, o cliente indica qual o endereço IP pretendido.

### ↳ 4º fase - Acknowledge

A mensagem DHCP Ack é encapsulada num pacote BootP Reply. O servidor identifica positivamente o aluguer do endereço fornecendo simultaneamente outras informações de interesse.

## Aluguer de Endereços

- T1 time (lease time) - tempo ao longo do qual o terminal deve tentar renovar o aluguer
- T2 time (5% " " ) - tempo ao longo do qual o terminal deve tentar renovar o aluguer se a 1º tentativa não for bem sucedida
- release time - tempo ao longo do qual o terminal deve deixar de usar o endereço IP se o aluguer não for renovado

## Existência de múltiplos servidores DHCP - Redundância e falhas de funcionamento

Requisito - gamos disjuntos de endereços nos diferentes servidores

## Outras mensagens:

DHCP Decline - o cliente rejeita a oferta e reinicia o processo de aluguer de endereços

DHCP Nach - o servidor informa que não pode satisfazer o pedido que foi feito através do DHCP Request

DHCP Release - o cliente informa que pretende terminar o aluguer

DHCP Inform - o cliente solicita apenas alguns parâmetros (p.e. endereço de DNS)

## Cliente e Servidor em Redes Diferentes

- Ativam nos roteadores → BootP Relay Agent
- Os roteadores encaminham os pacotes BootP Request para o endereço do servidor DHCP, colocando no campo Gateway IP Address o endereço da interface receptora
- O servidor envia os BootP Replys para o gateway IP Address

## Domain Name System - DNS

- Base de dados distribuída com serviço de tradução de nomes de estação (hostnames) em endereços IP, e vice-versa
- Organiza os nomes em domínios com uma estrutura hierárquica
- Cada sistema DNS define uma ou mais zonas sobre as quais tem autoridade de resolução

Hierarquia: domínio raiz  
domínio de topo TLD  
domínio de 2º nível  
domínio de 3º nível

## Name Server Records

- A (Address Records) - associa um nome a um endereço IPv4  
name.com IN A 4.29.81.184
- NS (Name Server Records) - define o servidor de nomes responsável pelo nome (pelo menos 2 endereços)  
AASEMI.ORG NS DPNS1.DNSNATSERV.ORG
- CNAME (Canonical name Records) - permite que uma máquina seja conhecida por mais de um nome  
mail.name.com IN CNAME nome.com
- SOA (Start of Authority Records) - determina que um endereço DNS é o ponto inicial da informação
- AAAA - igual ao Address Records mas para IPv6
- MX (mail exchange) - usado na troca de emails (SNMP)

## Domínios de Topo (TLD)

- gTLDs (generic TLDs) - os mais comuns .com, .edu, .gov...
- ccTLDs (country codes TLDs) - identificadores de um país .pt, .fr...
- cTLDs (corporate TLDs) - reservados para corporações internacionais
- pTLDs (public TLDs) - públicos .world, .air...
- dTLDs (deriving TLDs) - nomes que não são quaisquer nome associado

## Modelo de Gestão de Domínios

- Registry - entidade responsável por um determinado domínio (manutenção de zone file do TLD)
- Registrars - delegam das responsabilidades de gestão e comercialização dos domínios
- Registrar - pode definir um conjunto de resellers

- Reseller - vende os domínios em nome do registrante a troco de uma comissão
- Registrant - qualquer entidade que deseje registar um domínio

### Ciclo de vida de um domínio

um domínio pode ser registado por um período de 1 a 10 anos

Após esse período tem de ser renovado

No caso de não ser renovado é iniciado o processo de remoção do domínio da base de dados  
o domínio fica expirado 40 dias

se não for renovado vai para o período de 30 dias

No caso de ser pedido, ficará pendente e pode ser renovado

Esse é o ciclo quando não tem sido requerido período 5 dias é eliminado

### Informação de um registo

Nome servis, estado de um domínio, data de criação e expiração, contactos do registrante

### Resolução de nomes

As respostas recebidas são memorizadas temporariamente em cache de modo a serem apresentadas em futuros pedidos, melhorando assim a eficiência dos sistemas.

Resolução não recursiva - sucessivos pedidos e respostas para cada nível da hierarquia

Mais eficiente, praticamente não é usada

Resolução recursiva - apenas um único pedido ao LNS e os pedidos são feitos dentro da hierarquia  
Mais eficiente

### Switching

Hubs / Repetidores - interligam segmentos do mesmo hpo de LANs

↳ apenas o nível físico (camada OSI 1)

Switch / Bridges - interligam diferentes VLANs

Funções adicionais ↳ Store e forward - ao invés de mandarem o pacote para todos os portos podem apenas mandar para o porto da estação destino  
os portos podem operar a diferentes tipos de transmissão

↳ Com switches os colisões deixam de ser um problema. A interligação nível 2 evita o

### Aprendizagem de endereços - switches

Todos os switch têm uma tabela (tabela MAC) que mapeia os endereços dos dispositivos conectados às interfaces do switch.

Quando os switchs são ligados estes tabelas encontram-se vazias.

Quando um switch recebe uma trama MAC numa porta de entrada:

- regista na sua tabela a porta em que recebeu a trama e o endereço MAC de origem da trama
- Procura o endereço destino da trama na sua tabela para encaminhar a trama
- Se o endereço MAC destino da trama não existe na sua tabela, o switch envia a trama para todas as suas portas, menos pela que recebeu a trama [flooding], atropelando destino vai-lhe responder e dizer que o endereço é dele e o switch vai guardar o mac na sua tabela
- Se o endereço MAC destino da trama existe na sua tabela, o switch envia a trama apenas para a porta registada na sua tabela [forwarding]

### Ruter - tipo store-and-forward

Opera no nível de rede (camada OSI 3)

comuta com base nos endereços de nível 3

### LANS Virtuais - VLANs

(a interligação entre VLANs é feita através de um ruter)

Uma VLAN é uma rede local que agrupa um conjunto de máquinas de maneira lógica e não física.

↳ Pontos Inter-switch - torna possível encaminhar o tráfego de várias VLANs através da mesma interface.

↳ Pontos Tag - torna possível diferenciar a que VLAN pertence o pacote recebido, pois permitem associar a cada pacote o ID da VLAN que o enviu. Norma 802.1Q

### Spanning Tree

Protocolo de equipamento de rede que permite resolver problemas de loop em redes comutadas, permitindo caminhos alternativos e permite ainda desactivar caminhos. O algoritmo determina o caminho de menor custo entre cada segmento separado por bridges ou switches. Caso haja algum problema nesse caminho, o algoritmo recalcule entre os existentes o novo caminho.

Equações de Bellman Quando os custos das ligações são não negativos:

cumprimento do percurso mínimo = cumprimento do arco que une esse nó ao nó que

lhe segue no percurso mínimo + cumprimento do percurso mínimo desse nó para o nó de origem

### Algoritmo de Bellman-Ford distribuído e assíncrono

Soma-se os estimativas recebidas dos vizinhos ao custo da porta por onde receberam o anúncio

Guardam o menor valor

## Encaminhamento baseado em spanning tree

- É escolhido um switch como nó raiz / tronco
- Os outros switch usam o algoritmo Bellman-Ford para calcular o vizinho no percurso de custo mínimo
- As ligações compostas pelos percursos de custo mínimo de todos os switches para o nó raiz definem um árvore abrangente (spanning tree)
- As portas activas são as das ligações que compõem a árvore abrangente
- É necessário um critério para desempatar quando há múltiplos percursos de custo mínimo

## Conceptos Básicos de spanning tree

Switch ID - cada switch é identificado por um endereço: 2 bytes-prioridade; 6 bytes-endereço MAC

Switch Raiz - switch que está na raiz da spanning tree → Switch com menor ID

Path Cost - custo associado a cada porta do switch (cada porta configurada)

Bridge Designada - bridge que numa LAN é responsável pelo envio dos pacotes da LAN para raiz e v.v.

↳ Bridge raiz é a bridge designada em todas as LANs que está ligada

Ponta Designada - porta que numa LAN é responsável pelo envio de pacotes da LAN para a raiz e v.v. (uma das portas da bridge designada)

Ponta raiz - porta responsável pela transmissão (inclusão) de pacotes de/para a bridge raiz

- Cada bridge tem associado um custo para a raiz (root path cost)
  - = soma dos custos das portas que recebem os pacotes enviados pela raiz
- A ponta raiz é em cada bridge a porta que fornece o melhor percurso para a raiz
- A ponta designada é em cada LAN a porta que fornece o melhor percurso para a raiz
- Pontos Activos → ponta raiz + pontas designadas (em cada bridge)

Protocolo IEEE 802.1D → BPDU (Bridge Protocol Data Units)

Para construir e manter a spanning tree os bridges trocam mensagens especiais entre si: BPDU

Dois tipos: Configuration e Topology Change Notification

Configuração - configuração da spanning tree é feita pelos conf. BPDU

Campos importantes: Root ID, Root Path Cost, Bridge ID, Ponta ID → que envia mensagens

## Mantenimento da Spanning Tree

Periodicamente os bridges enviam pelos portas designadas Conf. BPDU = hello time

Periodicidade → hello time recomendado 2 segundos

Onderação das mensagens de configuração (ordenem dos campos)

valores todos menores	<u>Root ID</u>	<u>Root Path Cost</u>	<u>Bridge ID</u>	<u>Pont ID</u>
↓				
valores maiores				

## Construção da Spanning Tree

Cada bridge assume inicialmente que é a bridge raiz  $\rightarrow$  Root Path Cost = 0 envia mensagens de conf. p. todos os portes

Identifica a bridge Raiz

Identifica os root path cost e portos raiz

Identifica bridges designadas e portos designados

## Avaria nas bridges cujos VLAN's

Após a alteração da topologia da rede pode existir perda de conectividade se uma ponte que estava inativa aínde não se apercebeu que deve estar ativa na nova topologia e como inverso podem existir ciclos temporários.

Para minimizar a probabilidade de ciclos as bridges são obrigadas a esperar algum tempo antes de permitirem que um dos seus portos mude de estado, tempo em função do forward delay

## Estados dos portos

blocking - os processos de aprendizagem e expedição de pacotes estão inibidos

listening - os processos de aprendizagem e expedição de pacotes estão inibidos, transita para o estado learning após um tempo de permanência neste estado

learning - o processo de aprendizagem está activo mas o de expedição está inibido, transita para o estado forwarding após algum tempo de permanência neste estado

forwarding - é o estado activo, o processo de aprendizagem e expedição estão activos

disabled - os processos de aprendizagem e expedição estão inibidos, não participa no spanning tree

## tempo de vida das Entraidas da tabela de encaminhamento

2 tempos de largo: usado por defeito (valor recomendado = 5 minutos)

vida corta: usado quando o spanning tree está em recuperação (15 segundos) - exige processo de reabilitação de alterações da topologia da rede

tempo de vida demasiado longo - no exagerando de pacotes pendentes

tempo de vida demasiado curto - trazendo ruído exagerado devido ao flooding

## Outros protocolos

IEEE 802.1p - qualidade de serviço com base em prioridades

define o Campo User Priority (3 bits) que permitem 8 níveis de prioridade

+ prioridade para o tráfego crítico da gestão de rede

IEEE 802.1w - rapid spanning tree - acelera os tempos de convergência da ST no caso de alteração da topologia

IEEE 802.1s multiple Spanning tree - permite criar múltiplos spanning trees e atribuir a cada VLAN uma das spanning trees criadas. Permite criar múltiplas regiões

## Routing

Sistemas Autónomos - conjunto de roteadores com uma política de encaminhamento própria e responsabilidade de uma única administração

É identificado por um endereço único de 16 bits (pode ser atribuído por um ISP)

Protocolos de encaminhamento no interior: RIP, OSPF, IS-IS

Encaminhamento entre AS: BGP

Distance Vector - cada roteador conhece a informação que os roteadores vizinhos enviam periodicamente

Cada roteador determina os percursos de custo mínimo → baseado no algoritmo Bellman-Ford

Ex: RIP

Link-State - os roteadores conhecem a topologia completa da rede e usam um algoritmo baseado para determinar de custo mínimo para todos os destinos

A informação para construir a tabela cada roteador é obtida por flooding

Ex: OSPF

### RIP - Routing Information Protocol

- Protocolo tipo distância vetor
- Cada roteador tem um vetor distância constituída por uma lista de endereços IP que conhece e para cada qual estimativa de melhor custo. Cada roteador envia periodicamente o seu vetor distância para os seus vizinhos.
- O custo de percurso de um roteador para uma rede é dado pelo nº de roteadores intermédios
- Cada roteador determina as entradas da sua tabela com base nas distâncias recebidas dos vizinhos. Baseado em troca de mensagens entre roteadores que usam este protocolo. Cada mensagem trocada contém informação sobre as redes que o roteador conhece e a distância do roteador para cada uma das rotas. O roteador que recebe as mensagens, calcula a distância para os outros roteadores e guarda na sua tabela de encaminhamento. Esta distância é desigualada por hopped-1 de saltos (roteadores existentes no caminho). Mensagens trocadas em intervalos regulares.]

### Mensagens

RIP Request (apenas) - enviado por um roteador que foi enviado recentemente ligado ou quando a validade da informação atinge o um destino expira.

RIP Response - contém um vetor distância (com/sem split horizon) e é enviado periodicamente, especificamente (triggered updates) e em resposta a um RIP request

Com split horizon - cada roteador anuncia o vetor distância completo para todas as interfaces

Sem split horizon - em cada interface, o roteador anuncia apenas os roteadores destino para os quais essa interface não é usada no encaminhamento dos pacotes

→ o split horizon diminui o tempo de convergência das tabelas de encaminhamento

## OSPF (Open Shortest Path First)

Protocolo tipo link state

Cada roteador tem uma base de dados com a topologia da rede. A informação para a construção da base de dados é enviada através de flooding. Usa o algoritmo de dijkstra para calcular os percursos de custo mínimo. O custo de um percurso de um roteador para uma rede é dado pela soma dos custos das interfaces de saída na sentido do roteador para a rede.

!

### Tipos de Pacotes OSPF:

Hello: estabelecem relação de vizinhança entre vizinhos

Link State Description: enviam sumários de entradas link-state

Link State Request: pedem o conteúdo de uma entrada link-state

" " Update: envia o conteúdo de uma entrada link-state

" " Acknowledgment: confirma a receção de uma entrada link-state

### Eleição do Designated Router e do Backup Designated Router

O 1º roteador a ser ligado é o DR e o segundo é o BDR

se o DR falhar o BDR passa a ser DR e o novo BDR será o roteador com maior prioridade de todos os outros roteadores. Em caso de empate é escolhido o roteador com ID maior.

### Bases de Dados OSPF: identificação dos roteadores e das redes

OSPF Router ID - endereço IP de uma das suas interfaces (o maior no instante de activação)

OSPF Network ID - endereço IP da interface do seu Designated Router (DR)

→ Bases de dados organizados em duas tabelas

informação relativa de todos os roteadores → Roteiros Link States

informação relativa a todas as redes intermédias → Net Link States

Redes - Network ID - Roteiros - Router ID

De cada roteador se armazena informação relativa às várias redes que estão diretamente ligadas

identificação do roteador, nº de interfaces, rede intermédia, network id, endereço de interface, custo

Relativamente a cada rede intermédia é armazenada informação relativa aos vários roteadores que lhe estão diretamente ligados

Network ID, roteadores diretamente ligados (In)

### Áreas no OSPF

Além de tornar o protocolo mais escalável, para as grandes empresas conceito de área.

Assim um roteador interno conhece apenas a topologia da sua área. Um Área Border Router conhece a topologia das áreas a que está ligado. O encaminhamento entre as áreas tem de passar pela área de backbone.

### Parâmetros Hello São utilizados:

Para descobrir quem são os seus vizinhos em cada interface e para eleger o DR e o BDR em cada Rede. Parâmetros são enviados de 10 em 10 segundos e têm como objectivo detectar falhas de conectividade. Não transportam informação de encaminhamento.

### Diferenças entre RIP e OSPF

RIP: Simples, encaminhamento baseado em saltos, não é escalável, processamento contínuo de tab. encaminhamento.

OSPF: Complexo, escalável para grandes redes, processamento parcial das tabelas de encaminhamento, utilização intensa da rede apenas em processos de flooding, convergência de tabelas de encaminhamento mais rápida.

### BGP (Border Gateway Protocol)

Protocolo de encaminhamento de diferentes sistemas autónomos. Usa como protocolo de transporte o TCP e o número de porta 179. Número do SA definido por 4 bytes.

#### BGP interno e externo

↳ As relações de vizinhança podem ser estabelecidas internamente de um mesmo SA ou de diferentes SA.

Sistema Autônomo Single-homed - possui apenas um nó fronteira para atingir redes fora (domínio).

Sistema Autônomo multi-homed non-transit - possui mais do que um nó fronteira, não permite que o tráfego atravesse.

Sistema Autônomo transit multi-homed - possui mais do que um nó fronteira para a extensão e suporte tráfego de trânsito.

### Parâmetros BGP

Open - começam por estabelecer relações de vizinhança (p.e. declaram o nº de SA)

Inicialmente não trocados todos os not. BGP

Update - informaçõe de encaminhamento

Keepalive - quando não há alterações de notas, não enviaidas periodicamente entre vizinhos para manter relações de vizinhança

Notification - transmitidos para reportar situações de erro e terminar relações

#### ↳ Mensagens Update

Withdraw routes - lista de redes IP que já não podem ser atingidas

Path Attributes - permite implementar políticas de encaminhamento → AS-Path: SA no percurso para o destino

Network layer reachability information - listagem dos nós destino anunciamos

↳ Agregação BGP - Quando um nó fronteira divide 2 sub-redes de outro SA sempre Pela interface nega para terceiros com uma entrada na sua tabela de encaminhamento.

Atributos BGP - é uma métrica usada para descrever as características de um caminho BGP.

(atributos incluídos nas mensagens de update trocadas pelos peers BGP para anunciar rotas)

- well-known mandatory - AS-path, next-hop, Origin
- well-known discretionary - local preference, atomic Aggregate
- option transitive - Aggregator, Community, AS-Path, AS-Set, path
- option non-transitive - podem não ser suportados por todas as implementações BGP

↳

AS-path - quando o anúncio de uma rota passa através de um sistema Autônomo, o P do SA é adicionado a uma lista ordenada de SA que corresponde aos sistemas que o anúncio já atravessou.

Origin - indica como é que o BGP aprendeu a informação relativa a uma determinada rota.

next-hop - endereço usado para alcançar o roteador anunciante. Para o EBGP o next-hop é o endereço IP da ligação dos peers. Para o IBGP o endereço next-hop é transportado pelo SA local.

local preference - é usado para escolher um ponto de saída do SA local e é propagado para todo o SA local. Se houverem vários é usado para selecionar uma rota específica.

Atomic Aggregation - ajuda os roteadores que centos de rotas específicas foram agregadas para uma 'rota específica'. São perdidos mais rotas específicas.

Aggregator - fornece info sobre o SA que realizou a agregação e o end. IP de roteador que agregou.

Community - agrupa rotas que partilham propriedades comuns de tal forma que possam ser aplicadas políticas de agregado.

No-export - não anuncia esta rota aos peers EBGP.

No-advertise - não comunica esta rota a nenhum peer.

Internet - anuncia esta rota à comunidade, todos os roteadores da rede pertencem a ela.

weight - definido pelo circo, roteadores sabem tudo que é uma rota, a rota com weight eleito é a preferida.

### Seleção de Caminhos BGP

O BGP pode receber múltiplos anúncios para uma mesma rota.

Seleciona apenas um caminho como melhor.

Coloca o caminho selecionado na tabela de encaminhamento e propaga-o aos seus vizinhos.

Utiliza os critérios (par ordenem): weight, maior local preference, caminho originado localmente, shortest path.

Filtagem BGP e Route Flaps - envio/recepção de updates pode ser controlado por diferentes filtros.

- Filtagem com base: info de rota, informação de caminho, comunidades

- Route Flaps: controlar e modificar a informação de encaminhamento e definir condições de qual rota não redistribuir entre dominios e encaminhamento.

Sincronização BGP - se o seu SA encaminha tráfego de outro SA para um terceiro SA, o BGP

não deve anunciar a rota antes que todos os roteadores do seu SA tenham aprendido essa rota via IGP.

(espera até que o IGP tenha propagado a rota no interior do SA. Então anuncia a rota aos peers externos.)

## NAT e NAPT

\* NAT - faz a tradução entre endereços privados e públicos

NAPT - para além dos endereços faz a tradução entre nº de portos UDP ou TCP

Associações entre endereços públicos e privados podem ser estáticas ou dinâmicas

### • Estática

Correspondência entre o endereço NAT e o endereço público é contígua na rede estaticamente no neutro  
permite associações iniciadas nos dois sentidos

### Dinâmicas

A correspondência entre o endereço NAT e endereço público é feita automaticamente quando  
o primeiro pacote chega ao roteador NAT.

→ Para os endereços NAT só privados os pacotes para estes destinos não são encaminhados na rede pública  
funcionamento (associação dinâmica)

O roteador "fronteira" tem um conjunto de endereços por associação dinâmica

quando é enviado pacote de rede privada por a rede pública o endereço quando chega ao roteador fronteira  
é associado a um dos endereços disponíveis públicos e então é enviado por a seu destino público

As associações têm um tempo de vida associado

## NAPT

### Funcionamento

O roteador "fronteira" tem um conjunto de endereços públicos

contudo aqui é necessário traduzir o nº do porto origem porque os nº de porto são diferentes

a cada porta associar-se um endereço, assim todos os pacotes recebidos num mesmo porta são necessariamente  
para uma determinada máquina

regas de redirecionamento

## UDP e TCP

Endereço IP - identifica a ligação de uma estação a uma rede IP

nº de porta - identifica uma aplicação em execução numa estação

→ O SO assegura a atribuição de nº diferentes a cada aplicação

Alguns serviços fazem-se acompanhar de um nº de porta para o lado do servidor (Ex: 80 - HTTP)

## UDP - User Datagram Protocol

Proporciona um serviço de transporte de dados com as características de desempenho oferecidas pela  
rede IP. Permite a troca de dados entre aplicações e não apenas entre estações, através do cabeçalho com  
o identificador do nº de porta. Envio de dados para múltiplos destinos.

## Datagrama UDP

Source Port (octetos) - nº do porto da aplicação origem, é opcional, se não usada → preenchida com zeros

Destination Port (octetos) - nº do porto da aplicação destino

Datagram Length (n) - nº de octetos do datagrama (cabecalho + dados)

Checksum (octetos) - deteção de erros e validação dos extremos

Is calculado com base no datagrama completo IP + pseudoheshead → verifica se a mensagem foi enviada entre extremos conectados

### TCP - Transmission Control Protocol

Transporte de dados fiável, orientado à ligação, é bi-direcional, suporta apenas ligações ponto-a-ponto, faz uso da noção de fluxo de informação e proporciona o estabelecimento e a terminação da ligação transparente.

O TCP é uma sequência de protocolos, representa o conjunto de regras de comunicação e baseia-se na noção de endereçamento IP. Tem objectivos como o fracionamento de mensagens em pacotes, utilização de um sistema de endereços, encaminhamento de dados na rede e controlo de erros e transmissão de dados.]

Formato dum segmento TCP (significando quando estão a 1)

URG - campo urgent Point valido

ACK - campo acknowledgement valido

PSH - dados necessitam um push

RST - fazem reset à ligação

SYN - sincronizam sequence number

FIN - anigem terminou o envio

### Campos do Cabecalho TCP

hlen - indica o tamanho do cabecalho em múltiplos de octetos, quando há opções o campo padding acrescenta octetos ao cabecalho

Sequence Number - simula os dados já enviados

⚠ exemplo slides

Acknowledge Number - simula os dados já recebidos

Window - permite ao emissor informar o receptor quantos octetos está preparado para receber

→ Sequence number → fluxo de dados no sentido da transmissão

→ Acknowledge number, window → fluxo de dados no sentido contrário

### Controlo de congestão no TCP

- situações de congestão ou falta de conectividade da rede IP provocam almasas e apiles de pacotes
- TCP inclui algoritmos que permitem recorrer de forma eficiente a estas situações
- 2 tipos de algoritmos:
  - ↳ Gestão de tempo e retransmissão de pacotes
  - ↳ Gestão de janelas e transmissão de pacotes

## Gestão do Janelo de Transmissão do TCP

$cwnd = \min(\text{credit}, cwnd) \rightarrow$  janelo de congestionamento, segmentos

janelo permitido em

segmentos

credito não autorizado

pelo ACK, segmentos

Procedimento slow start  $\rightarrow$  demasiado agressivo em situações de recuperacão do timeout

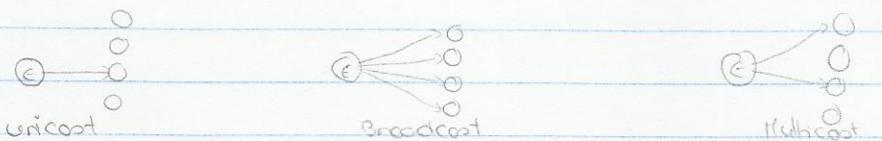
- $cwnd = \text{riss}$  (Maximum Segment Size)
- para que o atraso até ser atingido uma taxa de transmissão razoável não seja muito elevada o TCP aumenta a janela mais rapidamente nesta fase
- $cwnd$  cresce exponencialmente

Regras de gestão de janelas quando ocorre um timeout  $\rightarrow$  algoritmo de Jacobson

## Multicast

Refere-se a comunicações um para muitos ou muitos para muitos

- multicast ao nível da rede
- multicast ao nível da aplicação



A informação é enviada por uma aplicação originária com uma operação de envio e é recebida por múltiplas aplicações destino em diferentes estações.

Alternativa 1: a stack do protocolo TCP/IP da estação emissora estabelece ligações ponto a ponto com todos os destinos e envia múltiplos cópias uma por destino

Vantagens: redes sem capacidade multicast, utilização do TCP com todos os seus vantagens  
desvantagem: exige que a aplicação emissora especifique a lista de endereços destino

Alternativa 2: a estação envia cada pacote IP uma vez e a rede copia para múltiplos destinos

Desvantagens: capacidade de multicast, utilização UDP com todos os seus desvantagens

Vantagens: aplicação mais eficiente dos recursos da rede

- As Redes IP multicast não são redes connectless "como no caso das redes 'unicast'", é necessário estabelecer encaminhamentos multicast entre os roteadores para saberem como encaminhar pacotes multicast.
- Necessário: sincronização (entre estações e roteadores) e protocolos de encaminhamento (entre roteadores)

## Identificação das estações destino

Faz-se uso dos endereços IP de classe D (começado por 1110), os participantes combinam o uso de um endereço que identifica a sessão. As estações receptoras anunciam aos roteadores a participação na sessão multicast (denunciada pelo endereço combinado (IGMP))

Os roteadores encaminham os pacotes IP enviados com o endereço destino escolhido para todos os nós que hoje estão presentes na respectiva sessão.

Endereços Classe D - alguns endereços estão reservados pela IANA. A gama 239.0.0.0 a 239.255.255.255 destina-se às redes privadas.

### IIGP - Internet Group Membership Protocol

- Aberto entre cada estação e os roteadores que lhe estão directamente ligados
- Serve para a estação anunciar ao roteador que quer participar numa sessão multicast
- Começa sobre o protocolo IP
- Os pacotes são enviados para o endereço destino 224.0.0.1 ("All Hosts")

### Mensagens IIGP

GIG - General Membership Query - enviado pelos roteadores para perguntar se existe alguma estação multicast

SIG - Specific Membership Query - enviado pelos roteadores para perguntar se existe alguma estação multicast específica.

MR - Membership Request - enviado pelas estações para sinalizarem que participaram numa sessão multicast

LGR - Leave Group Report - enviado pelas estações para sinalizarem que deixam de participar numa sessão multicast

Query Router - roteador com menor endereço IP na interface ligada à rede é aquele que mantém o "diálogo" IIGP com os terminais.

- Os roteadores enviam periodicamente GIG especificando um Maximum Response Time (MRT)
- Cada estação espeta um tempo RAND entre o MRT para responder com um MR especificando um endereço MR
- Se entretanto uma sessão já tem um MR para a mesma sessão, abre-se o seu envio do MR
- Uma estação emite um MR quando quer pertencer a uma sessão multicast.
- Uma estação emite um LGR quando quer deixar uma sessão
- Quando o roteador recebe um LGR emite um SIG para verificar se ainda há estações pertencentes a essa sessão.

### Conclusões finais

Qualquer estação pode juntar-se a uma sessão multicast recebendo/enviando informações. A formação de sessões é iniciada pelos receptores → os emissores não têm poder que estações possam receber informações.

A rede não providencia filtragem, ordenação ou privacidade dos pacotes multicast.

Filosofia da Unicast: com protocolos simples e livres → funcionalidades adicionais → comodas supérfluas

### Encaminhamento Multicast

Group-based tree - determinação de uma única árvore de encaminhamento por cada sessão multicast que interligue todos os nós com estações pertencentes à sessão.

Source-based tree - baseia-se na determinação de uma árvore de encaminhamento por cada sessão multicast e por cada emissão.

### Group-based tree

Árvore de Steiner: determinação da árvore de custo mínimo que interliga os nós com estações de uma sessão

→ Protocolo tipo Link-State

→ Complexidade exponencial

Rendezvous Point - árvore de custos de custo mínimo para um nó central

→ Nó central previamente escolhido e conhecido dos outros nós

→ Para se juntarem à árvore os nós com estações terminais da sessão enviam uma mensagem "join" pelo encurso de custo mínimo entre eles e o nó central

### Source-based tree

Quando o emissor é conhecido

→ Cada nó com recepções interessadas numa sessão multicast de um determinado emissor envia uma mensagem  $\text{join}(E, R)$  em direção ao emissor pelo custo mínimo.

→ No encurso join cada nó recebe esta mensagem pela interface Id, reenvia pela interface Io e constrói a tabela de encaminhamento  $(E, R)$  para os pacotes multicast que entram por Io para serem encaminhados por Id.

→ Encaminhamento multicast → baseado no endereço destino e no endereço próximo

→ Múltiplos emissores → árvore de encaminhamento multicast por emissor

→ Quando o receptor deixa de estar interessado, o nó que lhe está ligado envia uma mensagem prune em direção ao emissor

→ Mensagem Prune é reenviada apenas pelos nós que deixam de pertencer à árvore de encaminhamento

Quando o emissor não é conhecido

Railene Path Forwarding - criação de uma árvore virtual abrangente

→ Cada nó N encaminha os pacotes do nó origem, recibidos pelo nó anterior ao nó N se este fizer parte anterior no encurso de custo mínimo

→ Um nó sem estações terminais interessadas na sessão e sem vizinhos envia uma mensagem prune para o vizinho para deixar de receber pacotes multicast

Railene Path Forwarding com Prune - apagamento de nós membro de uma sessão multicast

3º Estratégia - mensagem graft para anular um prune

Estratégia - associa um tempo de vida ao prune e ao fim do qual os pacotes voltam a ser encaminhados.

#### Group-based tree

Minimiza informação nos roteiros

Limita nº de ligações para tráfego multicast

Problemas de congestionamento → reduzido de ligações

#### Source-based tree

Perdida de informação nos roteiros

Distribui tráfego por mais de ligações

mais problemas de congestionamento

#### DV-TCP - Distance Vector Multicast Routing Protocol

- Tipo source-based tree → tipo Distance Vector (parecido com RIP)
- Usa estratégia RPF com "prunning"
  - ↳ Distâncias dadas em nº de saltos ("hops")
  - ↳ v0 → distância de cada possível origem
  - ↳ Para cada possível origem, o roteador anuncia também as suas vizinhas quando só o último salto no percurso desde a origem
- Mensagens "Prune" - enviadas com um tempo de vida
- Mensagens "Graft" - serve para eliminar uma mensagem "prune"

#### MOSPF (Multicast Open Shortest Path First)

Tipo source-based tree → RPF com pruning

Faz uso do facto de cada roteador conhecer a topologia completa da rede

cada roteador pode processar localmente à medida das percorridas de custo mínimo

não suporta túneis

#### PIM - Protocol-Independent Multicast

##### PIM Dense Mode

- A maioria das redes tem estações que pretendem usar encaminhamento multicast
- A maioria dos roteadores da rede necessita de encaminhar pacotes multicast

##### PIM Sparse Mode

- As estações que pretendem usar encaminhamento multicast concentram-se num número reduzido de nades
- O nº de roteadores que necessitam de encaminhar pacotes multicast é pequeno comparativamente ao nº total de roteadores

##### PIM Dense Mode

Utilizado em situações em que os membros de um determinado grupo se encontram distribuídos por todo a rede e obriga assim grande parte dos roteadores da rede a participar no encerramento de envoi de datagramas multicast

O PIM-DT utiliza a técnica flood and prune. Inicialmente os pacotes multicast são enviados para toda a rede e depois só "sintetizam" todos os roteadores da rede que não tenham terminais interessados nessa interação.

Isto é: manda para todos os roteadores pacotes multicast mas quem quer ouvir os outros só cortados.

Flooding Inicial - manda os pacotes multicast para todos os seus vizinhos. Quando um roteador recebe tráfego num interface que não fornece o percurso de custo mínimo, descarta os pacotes.

Mensagem Prune - o tráfego é encaminhado para todos os vizinhos que não tenham enviado mensagens prune.

Roteador que não tem clientes interessados

Roteador que tem recebido tráfego multicast em mais do que um interface

Quando um roteador deixa de ter clientes interessados numa sessão multicast envia uma mensagem prune pelo interface onde está a receber o respectivo tráfego

Mensagem Graft - para um roteador receberem o tráfego multicast envia uma mensagem graft para anular o prune enviado anteriormente.

Mensagem Join - se uma mensagem prune é enviada para um meio partilhado caso existam roteadores interessados naquela sessão multicast devem enviar uma mensagem join

Mensagem Assert - Caso exista mais de um roteador a receber tráfego de uma determinada sessão multicast para um meio partilhado estes devem decidir qual é o responsável desse envio.

Todos enviam uma mensagem assert com o respectivo prego (é exibido o menor custo ou endereço IP maior)

### PIM Sparse Mode

Aqui o número de roteadores com unidades multicast é pequeno comparativamente ao número de roteadores da rede, especialmente quando lidamos com grandes quantidades de tráfego comparativamente à largura de banda. Quando um roteador (DR) recebe datagramas de um terminal para ser distribuído pela rede, este encapsula em mensagens PIM de controlo e encaminhamento unicast para o rendezvous point (RP). O RP é responsável pela distribuição das mensagens para os destinatários pretendidos.

Isto é: quando adere a uma sessão multicast o equipamento que adere pode só receber pacotes multicast.

### Adesão à Group-Shared Tree

Os roteadores receptores enviam uma mensagem join em direção ao RP de modo a se juntarem à group-shared tree que tem a sua raiz no Rendezvous Point.

### Nova fonte de tráfego

Um roteador emissor envia os pacotes multicast encapsulados em pacotes unicast para o RP através de uma mensagem register.

O RP recupera os pacotes multicast e devolve o pacote pelo Rume establecida, no primeiro pacote recebido envia uma mensagem join para o nó de emissor. Assim o nó de emissor começa a enviar pacotes pelo Source-based tree. O RP envia um registre Stop.

#### Adesão à Source-Based Tree

Quando o bit rate agregado excede o dobro da taxa de link, um nó de emissor pode optar por aderir à Source-based tree. Envia um join em direção à fonte do fluxo multicast que passa por todos os nós até encontrar um que pertence à Source-based tree.)

Quando um pacote for recebido através do Rume, é enviado uma mensagem RP-bit Phine em direção ao RP.