

"Se um café pela manhā nāo te acordar, tente executar o seguinte em um servidor que está em produção: rm -rf --no-preserve-root /"

Redes, Internet e Web

Paulo Ricardo Lisboa de Almeida





A internet

"A internet de hoje é discutivelmente o maior sistema já criado pela humanidade, contendo centenas de milhares de computadores, links de comunicação e roteadores; com bilhões de usuários conectados [...] e uma variedade de coisas conectadas na internet, incluindo videogames, sistemas de vigilância, relógios, óculos, termostatos e carros [...].

[...]Será que há alguma esperança de entendermos como isso funciona?" (Kurose, Ross; 2013).

A internet

"A internet de hoje é discutivelmente o maior sistema já criado pela humanidade, contendo centenas de milhares de computadores, links de comunicação e roteadores; com bilhões de usuários conectados [...] e uma variedade de coisas conectadas na internet, incluindo videogames, sistemas de vigilância, relógios, óculos, termostatos e carros [...].

[...]Será que há alguma esperança de entendermos como isso funciona?[...]

[...]Sim!" (Kurose, Ross; 2013).

Atenção

O que você vai ver agora é uma versão **muito** simplificada sobre como as coisas funcionam.

Serve como uma introdução ao assunto.

No **mundo real** as coisas são **mais complicadas**.

Você pode pesquisar detalhes na bibliografia da disciplina.

Começando pelo... começo

Vamos conectar dois computadores.



Começando pelo... começo

Vamos conectar dois computadores.

Desejamos que os computadores se comuniquem.

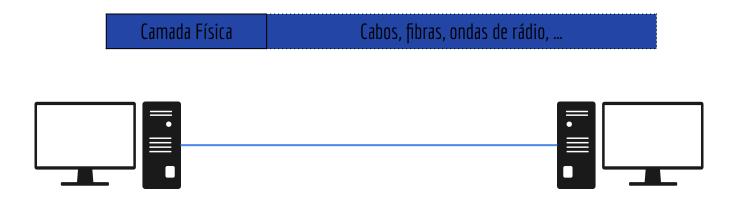
As informações passam por um cabo.

Poderia ser também por rádio, fibra óptica, sinal de fumaça, ...



Camada Física

As conexões físicas (cabos, fibras, ondas de rádio) entre os dispositivos são logicamente inseridos em uma camada chamada de camada física.



Pacotes





É necessário agora definir como os sinais são transmitidos. Qual o formato?

Pacotes





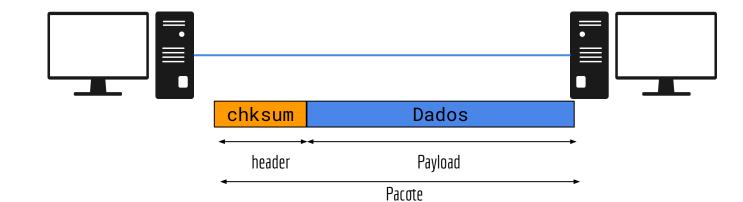
Dados
-
Payload
Pacote

Precisamos estabelecer protocolos de comunicação.

Um computador vai enviar um **pacote** (*packet*) para o outro.

Dentro do pacote, constam as informações que o computador deseja comunicar. **Payload** ou carga útil.

Pacotes



Precisamos estabelecer protocolos de comunicação.

Um computador vai enviar um **pacote** (*packet*) para o outro.

Dentro do pacote, constam as informações que o computador deseja comunicar.

Payload ou carga útil.

Informações de gerência do tráfego de dados.

Header ou cabeçalhos.

Exemplo: códigos de correção de erros (checksums).





Payload

Pacote

Cada pacote deve ter um tamanho máximo.

Exemplo: tipicamente 1.500 Bytes em uma conexão Ethernet.

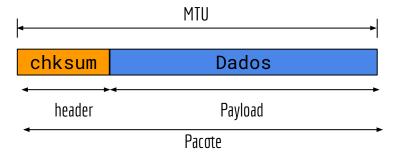
Esse tamanho máximo é chamado de MTU (Maximum Transfer Unit).

Se o dado a ser transmitido não couber em 1 MTU, precisamos quebrar a informação em múltiplos pacotes.

header

Enlace



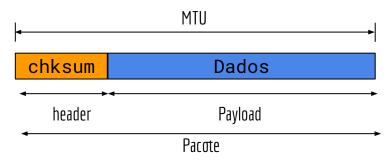


Inserir os headers e definir o MTU é tarefa para a camada de enlace.

Camada de enlace	MTU, definição de headers,
Camada Física	Cabos, fibras, ondas de rádio,







Dica do Linux:

Você pode configurar o tamanho do MTU do seu computadorou servidor com o comando ifconfig.

Veja como funciona verificando no manual: man ifconfig

Para verificar o MTU corrente: ifconfig | grep -i mtu

Cuidado: se você definir um MTU maior que o suportado pela rede:

Rede IPv4: Pode causar overhead nos roteadores, que precisarão quebrar os pacotes.

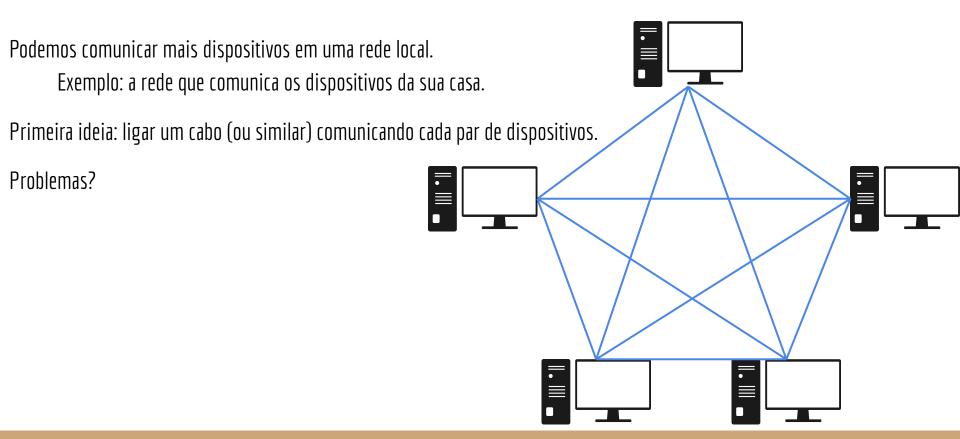
Rede IPv6: Retorna um erro.

Faça você mesmo

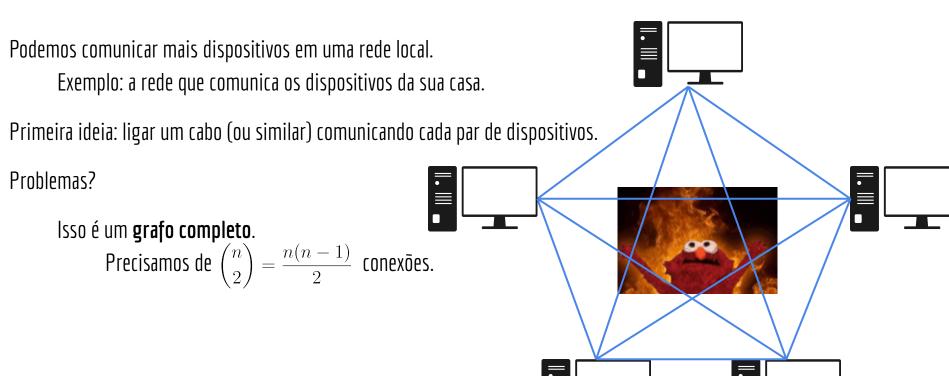
Verifique o MTU no computador que você está usando agora:

```
ifconfig | grep -i mtu
```

Aumentando a Rede

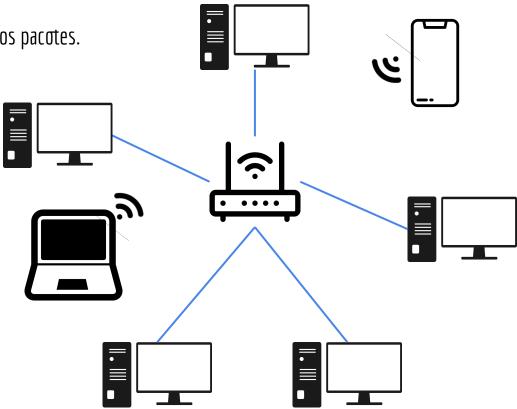


Aumentando a Rede



Podemos conectar a um dispositivo central que roteia os pacotes. Um **roteador** ou **router**.

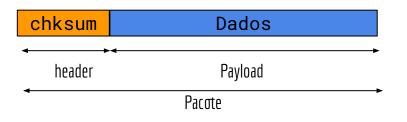
Obs.: em redes pequenas é comum o uso de um equipamento similar, mas menos "poderoso", chamado **switch.** Vamos usar um roteador nos exemplos para simplificar.

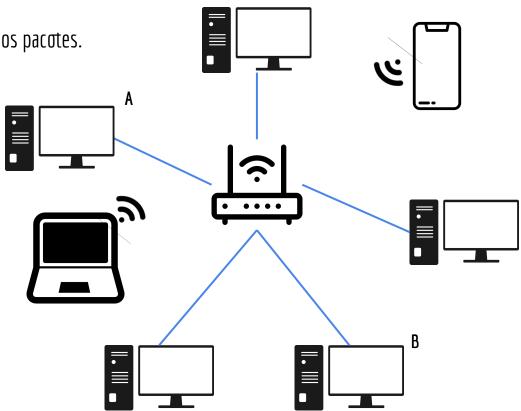


Podemos conectar a um dispositivo central que roteia os pacotes. Um **roteador** ou **router**.

Suponha que o computador A quer enviar um pacote para o computador B.

Qual o problema com nossos pacotes?



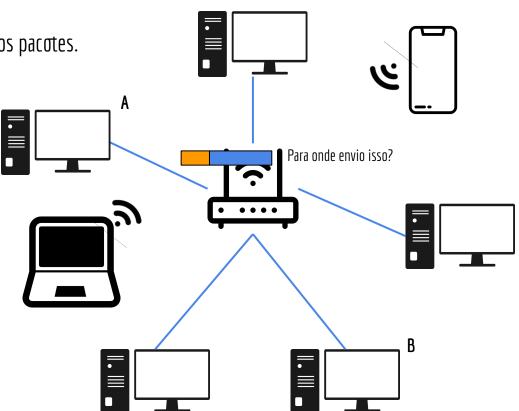


Podemos conectar a um dispositivo central que roteia os pacotes.

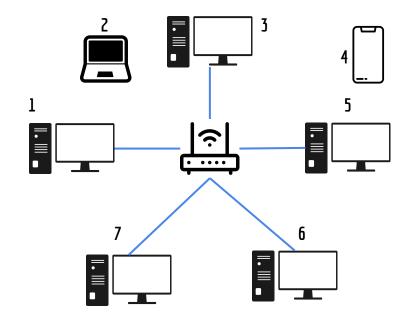
Um roteador ou router.

Suponha que o computador A quer enviar um pacote para o computador B.

Qual o problema com nossos pacotes?



Precisamos de um **endereço** interno para cada dispositivo.



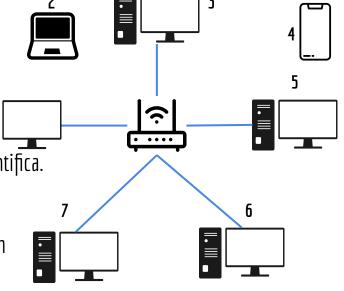
Precisamos de um **endereço** interno para cada dispositivo.

Os dispositivos de rede são feitos com um número "único" que os identifica. Endereço de hardware ou endereço físico.

Dispositivos comuns do nosso dia a dia, como Ethernet, Wifi e Bluetooth usam um endereço de 6 Bytes, comumente visto no formato hexadecimal.

Endereço MAC - *Media Access Control Address.*

XX:XX:XX:XX:XX

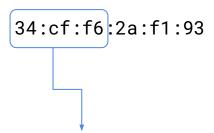


Você pode verificar o Endereço MAC dos seus dispositivos de rede digitando o comando *ifconfig* no Linux.

O endereço aparece como *ether* na saída do comando.

Você pode verificar o Endereço MAC dos seus dispositivos de rede digitando o comando *ifconfig* no Linux.

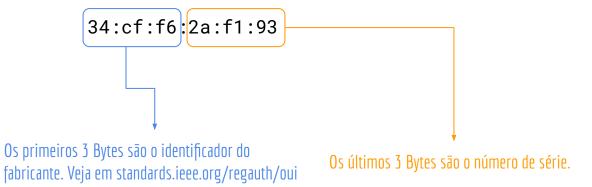
O endereço aparece como *ether* na saída do comando.



Os primeiros 3 Bytes são o identificador do fabricante. Veja em standards.ieee.org/regauth/oui

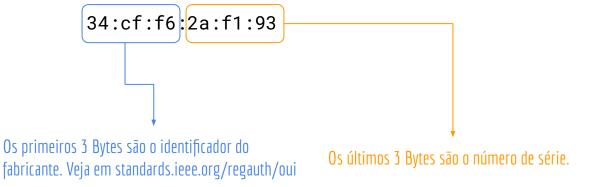
Você pode verificar o Endereço MAC dos seus dispositivos de rede digitando o comando *ifconfig* no Linux.

O endereço aparece como *ether* na saída do comando.



Você pode verificar o Endereço MAC dos seus dispositivos de rede digitando o comando *ifconfig* no Linux.

O endereço aparece como *ether* na saída do comando.



Faça você mesmo: identifique o fabricante do dispositivo de rede que você está usando agora.

Endereços MAC ficam na camada de enlace.

Endereços MAC são visíveis apenas na **rede local**. Um pacote não é transmitido com o endereço MAC na rede externa.

Na verdade, o IPv6 acaba adicionando o MAC como parte do IP.

Alguns hardwares permitem que o seu MAC seja trocado.

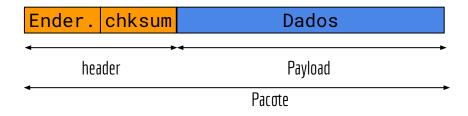
Especialmente útil em redes internas, quando precisamos substituir um hardware defeituoso por outro, mas precisamos manter o "endereço".

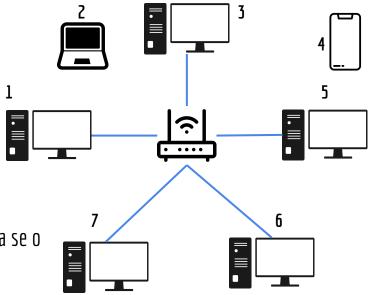
Camada de enlace	MTU, definição de headers,
Camada Física	Cabos, fibras, ondas de rádio,

O MAC pode ser adicionado no cabeçalho para identificar o destino.

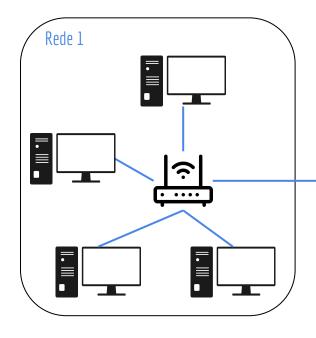
Opções:

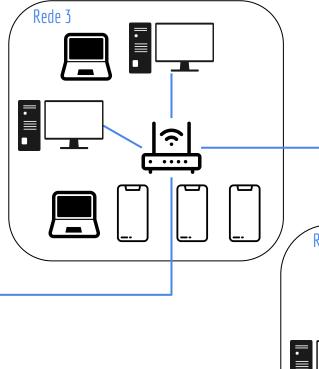
- 1. O roteador pode enviar o pacote para todos, e o destinatário verifica se o pacote deve ser lido ou não por ele.
- 2. O roteador pode enviar o pacote apenas para o destinatário.



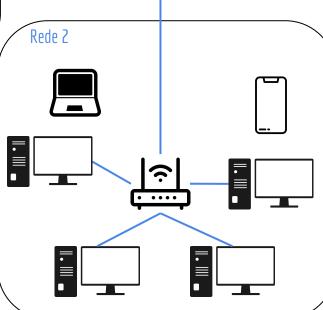


Indo além...





Podemos agora conectar múltiplas redes locais.

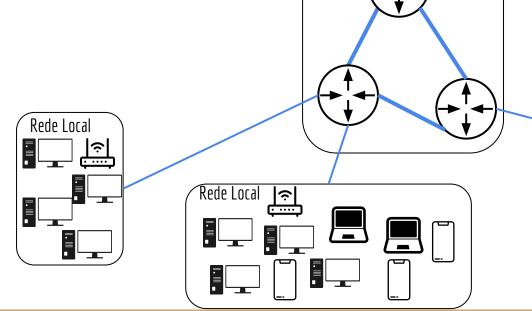


Indo além

Para evitar que as redes sejam conectadas diretamente umas às outras através de seus roteadores, podemos adicionar roteadores especiais, que interligam tudo em uma "via expressa".



Indo além



ISP



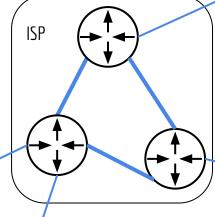




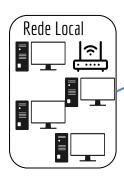
Indo além

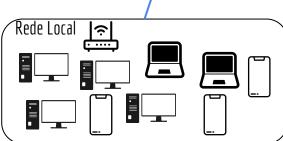
ISP: Internet service provider (ISP): Provê a conexão da sua rede local com as demais redes.

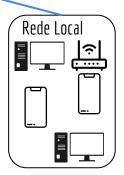
Contém seus próprios roteadores, que conversam com roteadores de outros ISPs.





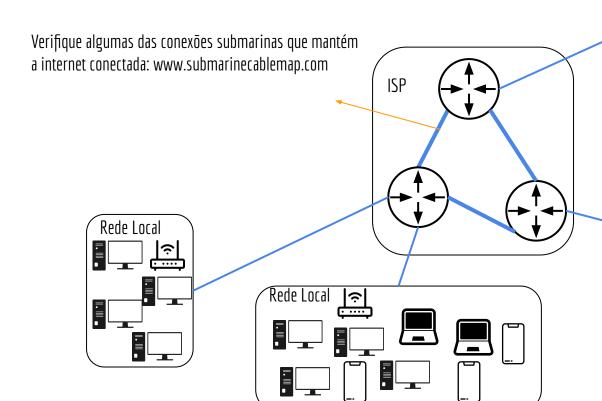








Verifique você mesmo







Veja esses vídeos



https://www.youtube.com/watch?v=d0gs497KApU



https://www.youtube.com/watch?v=bkgvqC3M1Tw

Conectando Tudo

Precisamos da mesma ideia de endereço, para rotear os pacotes.

Precisamos de um endereçamento mais robusto.

O que usamos é o IP - **Internet Protocol**.

Um número IP é um endereço independente de hardware.

Especifica um **destino único**.

1P

Podemos colocar o conceito de IP em uma camada extra de abstração, chamada de **rede**.

O mapeamento entre o endereço de IP e o hardware é feito na camada de enlace.

Exemplo: o seu roteador de casa pode fazer essa ligação.

Rede	IP
Camada de enlace	MTU, definição de headers,
Camada Física	Cabos, fibras, ondas de rádio,

IP e MAC

Uma rede IPv4 pode usar, por exemplo, o protocolo ARP para identificar o endereço IP associado a determinado MAC em uma rede local.

Rede	IP
Camada de enlace	MTU, definição de headers,
Camada Física	Cabos, fibras, ondas de rádio,

Um número IPv4 (que ainda é a maioria) é um número composto por 4 Bytes.

Você geralmente vai ver esse número em decimal na forma x.y.z.w, onde x,y,z e w são números entre [0-255].

Um número IPv4 (que ainda é a maioria) é um número composto por 4 Bytes.

Você geralmente vai ver esse número em decimal na forma x.y.z.w, onde x,y,z e w são números entre [0-255].

Idealmente, cada dispositivo conectado na internet deveria ter seu próprio IP.

Mas nem tudo são flores.

Com o IPv4, são possíveis apenas 2^{32} = 4.294.967.296 endereços únicos possíveis.

Um número IPv4 (que ainda é a maioria) é um número composto por 4 Bytes.

Você geralmente vai ver esse número em decimal na forma x.y.z.w, onde x,y,z e w são números entre [0-255].

Idealmente, cada dispositivo conectado na internet deveria ter seu próprio IP.

Mas nem tudo são flores.

Com o IPv4, são possíveis apenas 2^{32} = 4.294.967.296 endereços únicos possíveis.

O que geralmente se faz é ter um endereço IP "real" que referencia o roteador principal da sua rede local, e um endereço de IP "privado" para cada item da rede local, que só é usado internamente.

Uma das técnicas comuns para isso é o uso do DHCP - Veja na bibliografia.

IPv6

A versão melhorada do IPv4 é o IPv6.

É basicamente a mesma coisa, com algumas melhorias, e mais bits de endereço.

O IPv6 conta com 16 Bytes para endereços, ou seja, são possíveis $2^{128} = 3,402823669 \times 10^{38}$ endereços únicos possíveis.

Especialmente útil para Internet das Coisas - IoT.

IP_v6

A versão melhorada do IPv4 é o IPv6.

É basicamente a mesma coisa, com algumas melhorias, e mais bits de endereço.

O IPv6 conta com 16 Bytes para endereços, ou seja, são possíveis $2^{128} = 3,402823669 \times 10^{38}$ endereços únicos possíveis.

Especialmente útil para Internet das Coisas - IoT.

Onde cada x é um número hexadecimal [0,1,2,3,...,A,B,C,D,E,F].

IPv6

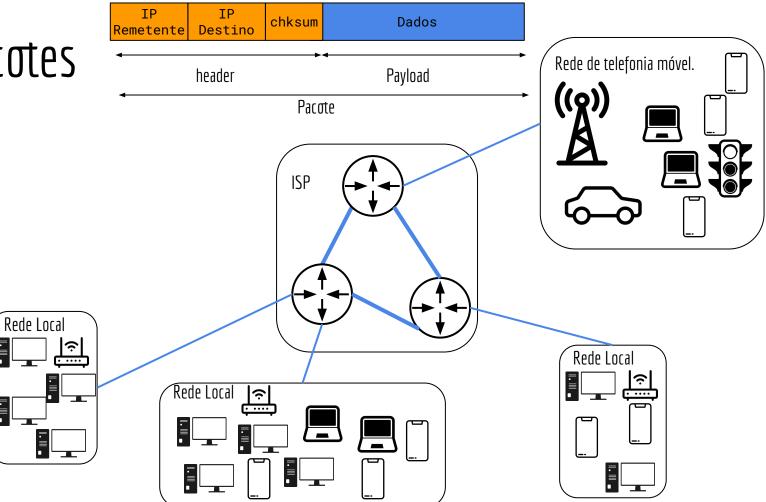
Os serviços estão migrando para IPv6, mas a um ritmo muito mais lento do que o esperado. Cerca de 40% dos acessos ao Google são feitos via dispositivos com endereço IPv6. Veja você mesmo: www.google.com/intl/en/ipv6/statistics.html

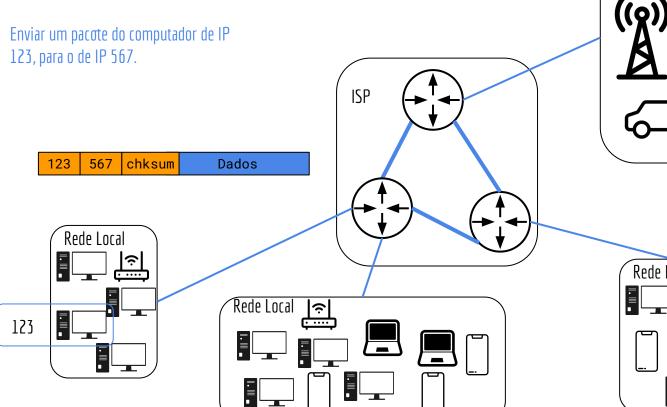
O número de IP é comumente dado pelo ISP.

Para simplificar, vamos considerar que cada dispositivo possui seu próprio endereço de IP. Sem nos preocupar se o IP é real ou só um IP interno.

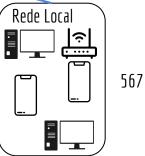
Você pode verificar o número de IP do seu computador com o comando ifconfig ou com ip address.

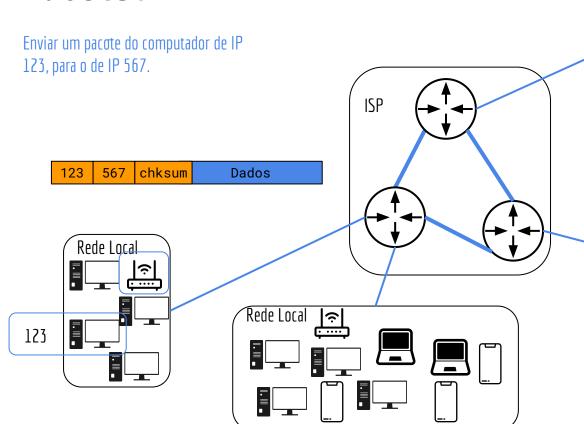
O número *IPv4* aparece como inet, e *IPv6* como inet6.



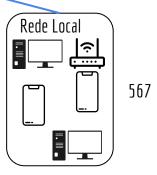


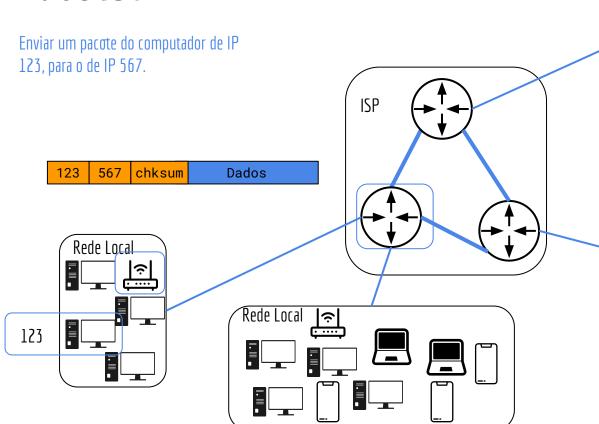




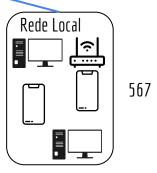


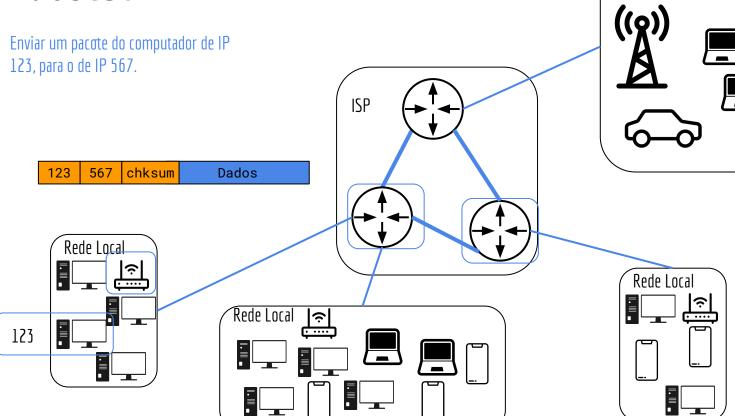






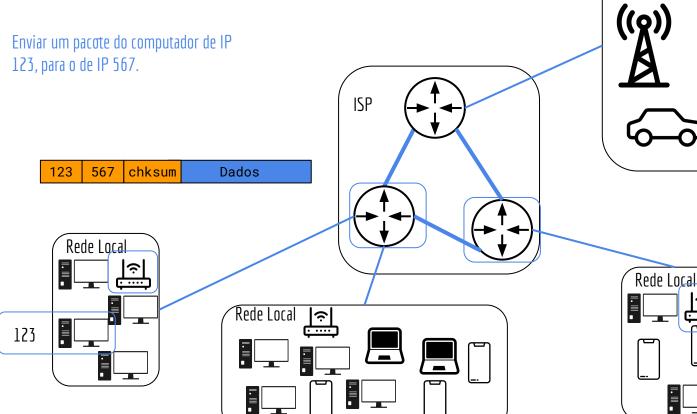




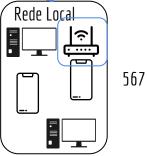


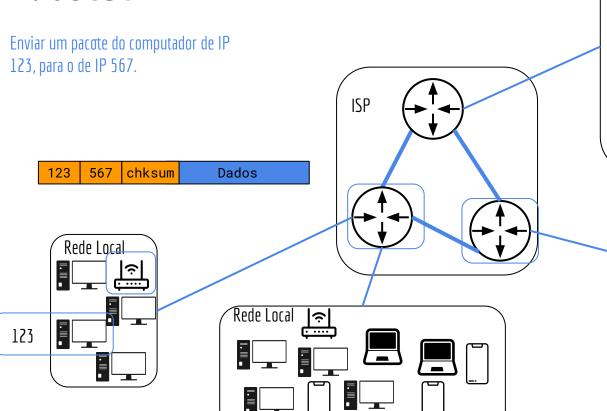
Rede de telefonia móvel.

567

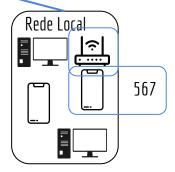












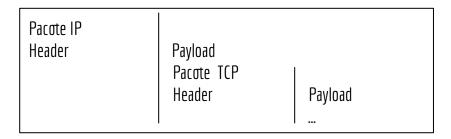
TCP/IP

A combinação dos protocolos de transmissão com o endereçamento IP dá o nome ao conjunto TCP/IP.

TCP: Transmission Control Protocol

A forma com que os pacotes são criados e endereçados é mais complexa do que a mostrada na aula. Mas o visto em aula serve como uma boa base.

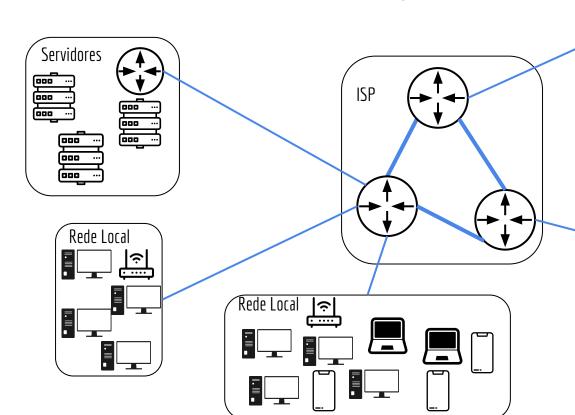
Na vida real, o que se faz é aninhar pacotes, um dentro do outro.



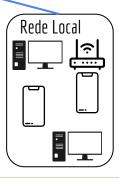
TCP/IP

Transporte	TCP,UDP.
Rede	IP
Camada de enlace	MTU, definição de headers,
Camada Física	Cabos, fibras, ondas de rádio,

Adicionando mais componentes



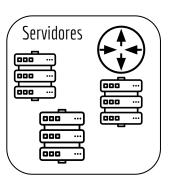




Servidor

Servidor: Um servidor é um "computador comum" que armazena e transmite informações, como páginas web, serviços de e-mail, dados de vídeo, ...

Você literalmente pode usar um computador comum, mas...



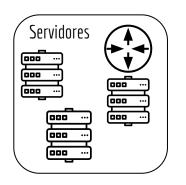
Servidor

Servidor: Um servidor é um "computador comum" que armazena e transmite informações, como páginas web, serviços de e-mail, dados de vídeo, ...

Você literalmente pode usar um computador comum, mas...

Idealmente você terá um computador com requisitos específicos, como redundância e capacidade para tratar diversas conexões simultaneamente.

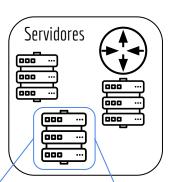
O seu e-mail (do Gmail, Microsoft, UFPR, ...) está armazenado em um servidor, em algum lugar.



Servidor

Quando você acessa o Google por exemplo, está acessando uma página Web, que está armazenada em algum servidor (computador).

O servidor recebe sua requisição, e envia os pacotes da página para o seu computador.





Falta uma peça nesse quebra-cabeças.

Para acessar o Google, precisamos interagir com o servidor onde o Google está hospedado.

Falta uma peça nesse quebra-cabeças.

Para acessar o Google, precisamos interagir com o servidor onde o Google está hospedado.

O servidor possui um endereço IP.

Mas não digitamos esse endereço para acessar o Google.

O endereço de IP existe, e de fato precisamos dele para acessar.

Para saber o endereço IP do Google, ou de qualquer página, basta usar o comando dig <endereço da página> no Linux.

O endereço vai aparecer em ANSWER SECTION.

O endereço de IP existe, e de fato precisamos dele para acessar.

Para saber o endereço IP do Google, ou de qualquer página, basta usar o comando dig <endereço da página> no Linux.

O endereço vai aparecer em *ANSWER SECTION*.

Faça você mesmo:

Use o comando dig para encontrar o endereço IP de páginas da internet, e acesse as páginas via o endereço IP diretamente no seu navegador.

Dica: não vai funcionar para qualquer página, já que algumas páginas estão em servidores compartilhados.

Seria incômodo acessar as páginas web através de seus endereços IP.

Da mesma forma que seria incômodo (mas talvez melhor) referenciar o endereço da sua casa por sua latitude e longitude.

Seria incômodo acessar as páginas web através de seus endereços IP.

Da mesma forma que seria incômodo (mas talvez melhor) referenciar o endereço da sua casa por sua latitude e longitude.

Precisamos mapear os endereços "humanos" (e.g., www.google.com) para os endereços IP dos servidores.

Quando acessamos uma página Web, o navegador autoMagicamente consegue inferir o endereço IP do servidor onde a página está hospedada.

Essa mágica é feita pelo resolvedor **DNS**.

Domain Name System (Sistema de Nomes de Domínio).

Resolvedores DNS

Os resolvedores DNS são bancos de dados descentralizados.

A sua rede local pode guardar um cache dos IPs de sites acessados recentemente.

Uma cache é uma pequena memória contendo algumas entradas.

Essa cache expira em determinado período de tempo (e.g., 1 dia)

Resolvedores DNS

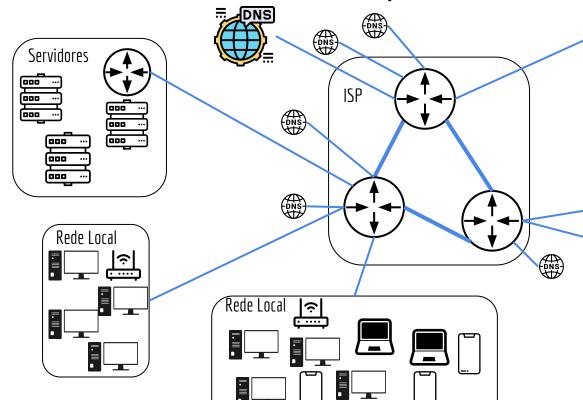
Os resolvedores DNS são bancos de dados descentralizados.

A sua rede local pode guardar um cache dos IPs de sites acessados recentemente.

Se a rede local não sabe o IP de um site, ela pode perguntar para, por exemplo, o seu ISP, que é outro cache.

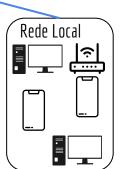
Se o ISP não sabe, pode perguntar, por exemplo, para um servidor raiz.

Adicionando mais componentes













Cache poisoning

O sistema DNS é um dos elos fracos da internet.

Comprometer um servidor (root ou cache) é um problema sério.

Um problema clássico que pode ocorrer em redes locais (como o da sua casa) é chamado de *cache poisoning*.

Cache poisoning



123.456.789.123

Meu Banco



222.222.222



Cache armazenada em seu roteador.



Um atacante teve acesso a cache, e modificou o IP do banco para um IP falso. O IP falso pode redirecionar para uma página falsa.

Faça você mesmo

Use o comando traceroute do Linux para verificar os endereços IP de todos os roteadores que você precisa passar para chegar a determinado endereço.

Exemplo:

traceroute www.google.com

Você pode tentar pesquisar a localização geográfica dos roteadores ou do endereço final em check-host.net.

Latência

A latência é o tempo que um pacote demora para chegar, por exemplo, de um servidor até sua máquina.

Uma latência grande pode ser devida a:

Uma conexão fisicamente distante entre as máquinas.

Exemplo: seu computador em Curitiba, e o servidor na Alemanha.

Passagem por muitos roteadores.

Cada roteador que recebe um pacote gasta tempo para roteá-lo.

Roteadores Sobrecarregados.

• ..

Latência

Como a latência pode nos prejudicar?

Latência

Como a latência pode nos prejudicar?

Você pode perder aquela partida no seu jogo favorito.



Latência

Como a latência pode nos prejudicar?

- Você pode perder aquela partida no seu jogo favorito.
- Em um caso mais sério (o jogo não era sério!?), isso inviabiliza que o processamento de determinadas aplicações seja feito remotamente.
 - Exemplo: um carro enviar as informações de piloto automático pela rede, para serem processadas em um servidor, é inviável.



Edge Computing

Muitos serviços têm migrado para edge computing por conta da latência e da largura de banda (capacidade) de redes.

O conceito básico é processar as informações diretamente na ponta (edge), sem enviar os dados crus para um servidor.

Exemplo:

Uma câmera pode processar as informações de tráfego sozinha, e enviar somente o número de carros na via no momento.

Não precisa enviar imagens para o servidor.

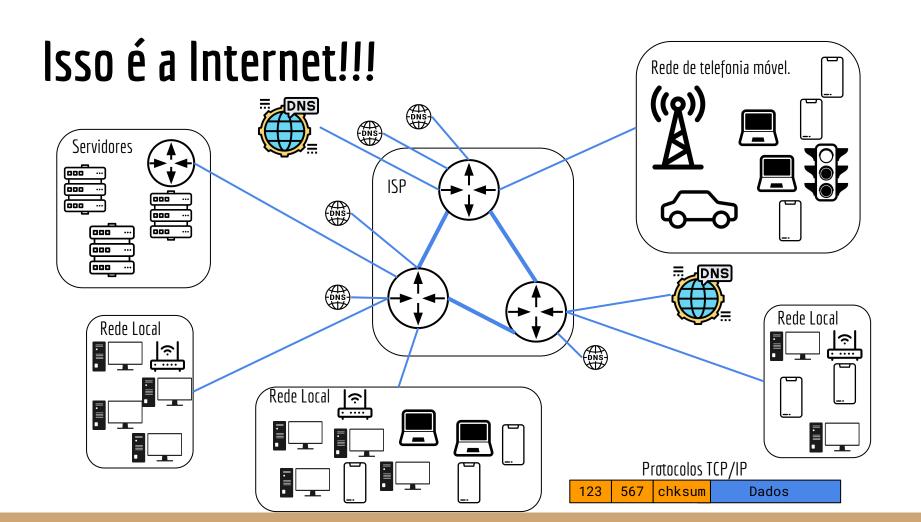


Latência

O comando clássico para medir a latência é o ping.

Exemplo: ping www.google.com

Use Control+c para parar.

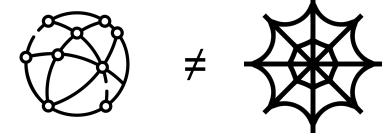


E a Web?

World Wide Web (Rede Mundial de Computadores): WWW ou **Web**

O nome não é dos melhores. Confunde a Web com a internet.

Web e internet **não** são a mesma coisa.



Web

A Web é uma coleção de recursos (documentos, páginas, ...) conectadas. Hiperlinks e outros identificadores.

Depende de protocolos específicos de transmissão, que ficam dentro da camada de aplicação. Ex.: protocolos HTTP e HTTPS.

Definem como informações (e.g., de páginas web) são transmitidas pela rede.

Aplicação	HTTP, HTTPS.
Transporte	TCP,UDP.
Rede	IP
Camada de enlace	ARP, MTU, definição de headers,
Camada Física	Cabos, fibras, ondas de rádio,





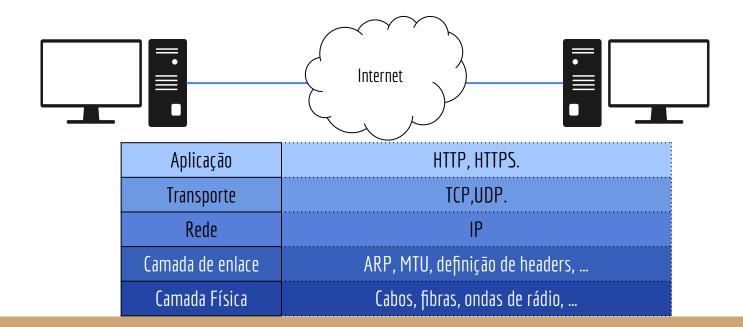




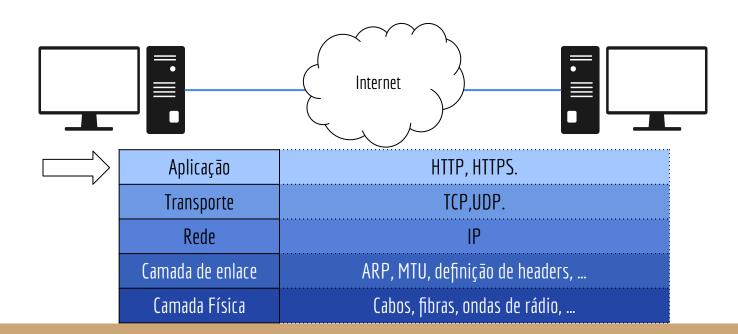


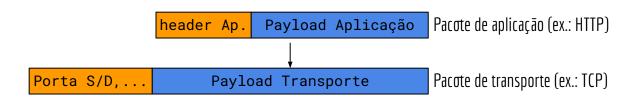
Como as coisas são realmente

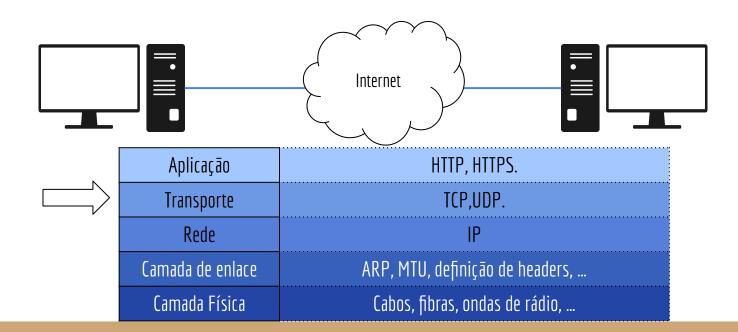
No mundo real, cada camada adiciona as informações necessárias para transmitir os dados.

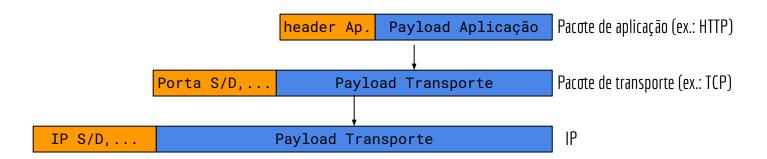


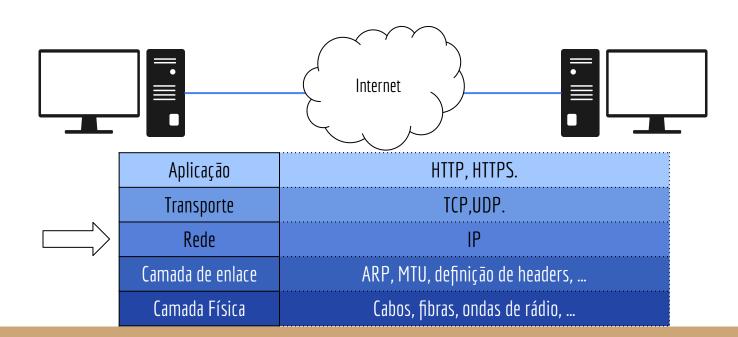


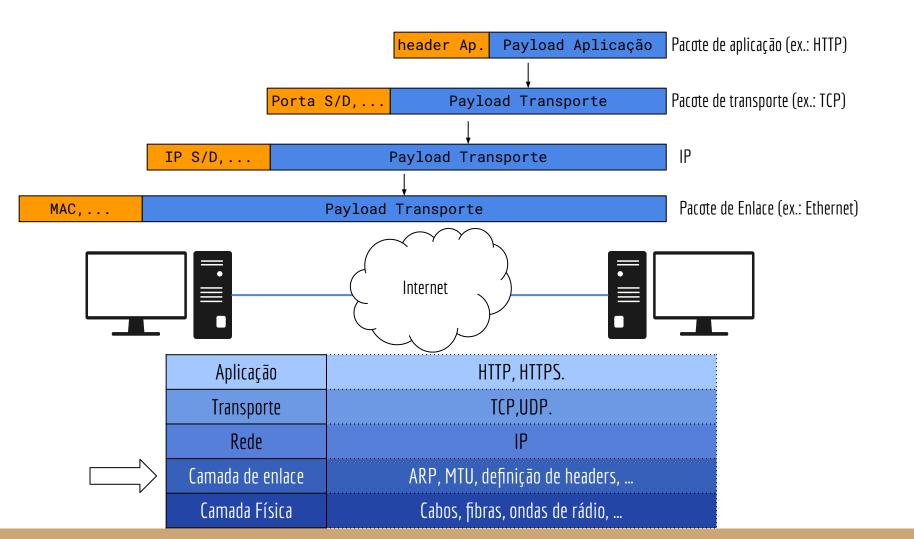


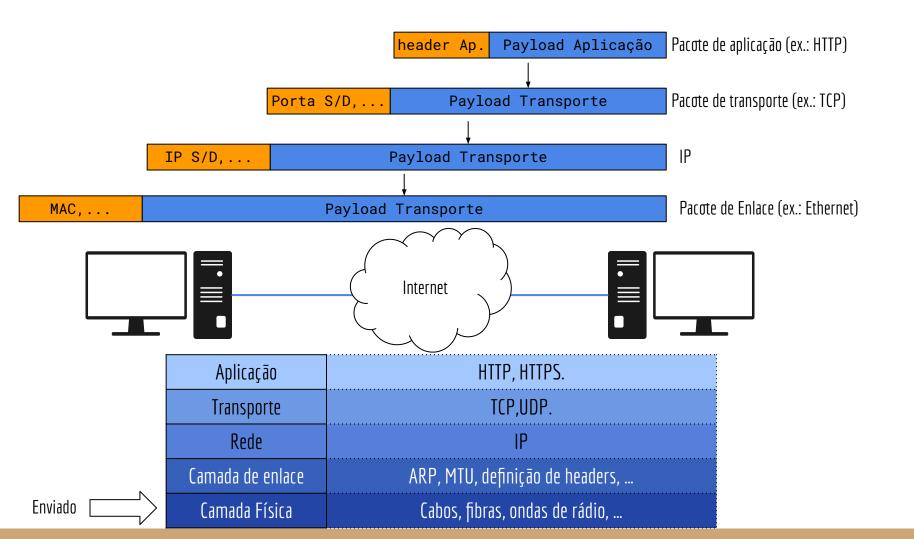












Por curiosidade

As páginas Web geralmente são criadas através de linguagens de marcação – HTML.

Muitas páginas precisam ser programadas (via linguagens de programação) para gerar o HTML dinamicamente.

Comum o uso de linguagens como:

PHP

Ruby

Python

Javascript

Go

•••

Por curiosidade

Você pode ver o html de qualquer página que você está visualizando.

No Firefox, por exemplo, basta digitar *Control+Shift+i* para inspecionar elementos específicos ou, *Control+U* para ver o código completo.

As camadas apresentadas são apenas uma uma simplificação.

As camadas clássicas são as do Modelo OSI.

A ideia é similar a apresentada na aula.

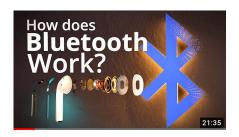
Veja na bibliografia.

Você também pode verificar na bibliografia como os pacotes são encapsulados, e todos os seus campos de *header*.

Veja esses vídeos:



Um pouco mais sobre a pilha de abstrações e protocolos da internet: youtu.be/PG9oKZdFb7w



Vídeo sobre redes Bluetooth: youtu.be/111vxu5qIUM



Como funciona o DNS: youtu.be/uOfonONtluk



Como funciona o NAT: youtu.be/01ajHxPLxAw



Mais formas de cache poisoning: youtu.be/7MT1F003_Yw

Exercícios

Instale o Wireshark no seu computador de casa.

https://www.wireshark.org

No linux basta usar sudo apt-get install wireshark pará instalar.

Depois de instalado, você precisa executar o wireshark com sudo: *sudo wireshark*, ou para executar sem sudo:

sudo dpkg-reconfigure wireshark-common sudo usermod -a -G wireshark \$USER Depois disso, reinicie sua sessão (ou reinicie o computador)

Inicie o wireshark, selecione a placa de rede que seu computador está conectado clicando duas vezes nela (na imagem de exemplo, *wlo1* é a interface wifi que a máquina está conectada. O wireshark captura todos os pacotes que chegam na sua rede. Você pode clicar em cada um, e ver o que tem dentro dos pacotes. Por exemplo, acesse um site, e veja os pacotes que chegam na sua máquina por conta desse acesso. Você pode inspecionar os headers e o payload de cada pacote.

2. Resolva os exercícios disponíveis no Moodle.

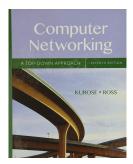


Bibliografia

Snyder et al. UNIX and Linux System Administration Handbook. 5a ed. 2017.



Kurose, , Ross. Redes de computadores e a internet: uma abordagem top-down. 2013.



Tanenbaum, Bos. Sistemas operacionais modernos. 4a ed. 2016.



Licença

Esta obra está licenciada com uma Licença <u>Creative Commons Atribuição 4.0 Internacional.</u>

