

Introdução à *internet*

Introdução

- Entender como funciona a *internet* é importante para o desenvolvimento de aplicações *web*.
- Apesar de muitas ferramentas automatizarem tarefas, a correta compreensão de como a informação trafega na *internet* pode auxiliar a resolver problemas que necessitam de algum ajuste fino.

Introdução

- A *internet* é simplesmente a intercomunicação de diferentes redes de computadores.
- As diversas redes comunicam-se através de grandes roteadores que ligam o planeta quase todo.

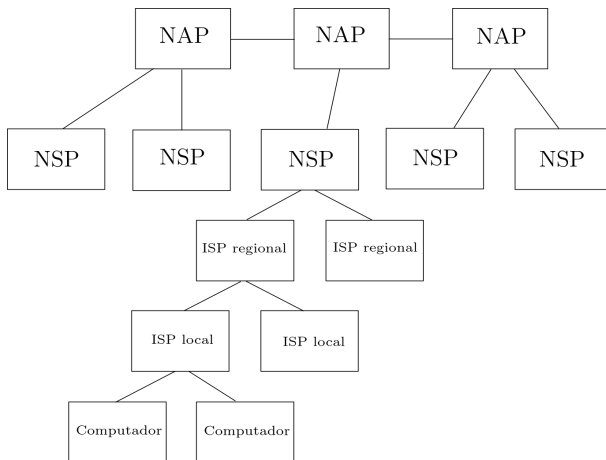
Introdução

- Como todo sistema computacional, é comum o foco apenas na abstração e o completo esquecimento de que a *internet* funciona sobre uma base física.
- Os dados nada mais são que sinais que trafegam em fios.
- Dada esta natureza física, é importante ter sempre em mente as limitações dos materiais.

Infraestrutura da *internet*

- A *internet*, assim como as demais redes de computadores, estão organizadas de maneira hierárquica.
- Há grandes NSP (*Network Service Providers*) interligados por NAPs (*Network Access Points*).
- Os NSPs também estão ligados a ISPs regionais que, por sua vez, estão ligados a ISPs locais.
- Os dispositivos de pessoas comuns e empresas estão ligados aos ISPs locais.
- A informação que sai de um computador vai a um ISP local, passa a um ISP regional, depois a NSP e, se necessário ainda, pode ir a outro NSP através de um NAP.

Infraestrutura da *internet*



Infraestrutura da *internet*

- Como é impossível que todos os dispositivos saibam onde estão localizados todos os demais dispositivos conectados à internet, a ligação entre os diferentes *backbones*¹ dos provedores é realizada através de roteadores.
- Cada roteador possui uma tabela de roteamento que permite escolher o melhor caminho para enviar um pacote de informação.
- Os dados trafegam por linhas físicas e que podem estar congestionadas. Então, o roteador deve escolher o caminho mais rápido.

¹Redes de comunicação de alta velocidade.

Protocolos

- Dada a quantidade de diferentes dispositivos e tipos de redes de computadores, foi necessário estabelecer protocolos de comunicação.
- Um protocolo de comunicação é uma série de regras que estabelece parâmetros e fluxos de comunicação.
- Através de um protocolo é possível a comunicação entre redes e dispositivos dos mais diversos fabricantes.

Protocolos

- Um protocolo único, porém, não é desejado, já que é difícil estabelecer uma regra universal.
- Apesar da *internet* como rede estar bem padronizada, as aplicações que rodam sobre esta infraestrutura evoluem rapidamente e necessitam de protocolos sempre mais atualizados que sejam condizentes com suas necessidades.

Protocolos

- A *internet* e suas aplicações funcionam com uma pilha de protocolos.
- Cada protocolo age em um determinado aspecto.
- Os aqui chamados aspectos são tecnicamente chamados de camadas.
- Cada camada tem suas responsabilidades e protocolos.
- Há protocolos para gerir o funcionamento das aplicações, do transporte de informações e da organização da rede.

Internet Protocol (IP)

- O protocolo que organiza os dispositivos na *internet* (camada de internet) é o *Internet Protocol* (IP).
- Ao conectar-se à internet através de um ISP, o dispositivo recebe um endereço IP temporário.
- Os servidores possuem endereços fixos.
- O endereço IP é formado por uma sequência numérica que tem formato variado de acordo com a versão do IP adotada.

Domain Name Service (DNS)

- Como os endereços IP são sequências numéricas de difícil memorização por seres humanos, criaram-se domínios.
- Os domínios são pseudônimos que permitem que os usuários possam digitar palavras ao invés de ter de decorar os endereços IPs.
- Exemplos de domínio são google.com, yahoo.com, etc.

Domain Name Service (DNS)

- Após o usuário solicitar um endereço de domínio no navegador, o navegador busca o endereço IP deste domínio através do *Domain Name Service* (DNS).
- O DNS é composto por uma série de servidores com tabelas indicando o endereço IP de cada domínio.
- Assim que recupera o endereço IP do domínio desejado, o navegador prossegue na tentativa de estabelecer conexão com o servidor localizado no dado endereço.

Transmission Control Protocol (TCP)

- O estabelecimento da conexão entre clientes e servidores é feito na camada de transporte.
- O protocolo da camada de transporte mais comum é o *Transmission Control Protocol* (TCP).
- Junto com o IP, o TCP é responsável por uma grande fatia de tráfego na *internet*.

Transmission Control Protocol (TCP)

- O TCP é responsável por encapsular a mensagem em pacotes que serão transmitidos pacotes, além de estabelecer as regras da conexão entre os dispositivos.
- O TCP é considerado confiável, pois ele estabelece que todos os pacotes devem ser entregues.
- Em caso de perda, o pacote será reenviado.
- A confiabilidade torna, porém, a comunicação mais lenta, pois assim que o pacote é recebido, deve-se enviar uma mensagem confirmando seu recebimento.
-

User Datagram Protocol (UDP)

- No caso de aplicações em que a perda de dados é tolerada, tem-se o UDP (*User Datagram Protocol*).
- No UDP, não há confirmação de recebimento e, por conseguinte, reenvio de pacotes perdidos.
- Este protocolo é menos confiável, mas muito mais rápido, sendo o ideal para aplicações em tempo real em que a perda de pacotes não seja sentida, como jogos, transmissões ao vivo, etc.

Camada de transporte

- Se a informação a ser enviada for muito grande, o protocolo de transporte (TCP ou UDP) pode quebrar a informação em diversos pacotes numerados e organizados em uma fila.
- Os pacotes são endereçados a uma porta específica do dispositivo receptor. As aplicações no dispositivo receptor devem "ouvir" a uma porta específica.

Camada de aplicação

- As camadas de transporte e redes não são o suficiente para rodar as aplicações na *internet*.
- Elas definem apenas a organização da rede e o modo de comunicação entre dois dispositivos.
- A maneira como as aplicações sobre esta estrutura irão funcionar é definida pelos protocolos de aplicação.
- O protocolo de aplicação varia de acordo com o tipo de aplicação (hipertexto, e-mail, transferência de arquivos, etc.)

Hypertext Transfer Protocol (HTTP)

- Para utilizar a *World Wide Web* (WWW), o protocolo utilizado é o *Hypertext Transfer Protocol* (HTTP).
- Ele define como as mensagens em hipertexto serão transferidas entre os dispositivos.

Hypertext Transfer Protocol (HTTP)

- No HTTP há duas entidades, clientes e servidores.
- Toda conexão HTTP deve ser iniciada pelo cliente, apesar de que, com o passar do tempo, tenham sido criados mecanismos que simulam mensagens iniciadas pelo servidor).
- O servidor recebe essa requisição e envia uma resposta.
- Por exemplo, quando é digitado um endereço em um navegador *web*, o que se está fazendo é uma requisição a um servidor.
- O servidor processa essa requisição e devolve um hipertexto a ser exibido no navegador cliente.

Hypertext Transfer Protocol (HTTP)

- O HTTP não controla o transporte dos pacotes, mas exige que as mensagens sejam enviadas através de um protocolo confiável, como o TCP. Antes de iniciar uma conexão HTTP, é necessário estabelecer uma conexão TCP.
- Para o HTTP, cada requisição é uma conexão independente de suas anteriores (*stateless*).
- O uso de *cookies* permite *stateful sessions*.

Hypertext Transfer Protocol (HTTP)

A requisição HTTP tem como elementos principais:

- método;
- endereço IP e porta;
- versão do protocolo;
- outros dados opcionais.

Hypertext Transfer Protocol (HTTP)

Os principais métodos HTTP são:

- GET → para requisições simples;
- POST → para enviar valores ao servidor;
- PUT → similar ao POST, porém idempotente;
- DELETE → para excluir recursos.

Hypertext Transfer Protocol (HTTP)

A resposta HTTP tem como elementos principais:

- versão do protocolo;
- código de *status*;
- mensagem de *status*;
- cabeçalhos;
- resposta.

Fluxo da *internet*

Em resumo, a *internet* tem o seguinte fluxo básico:

- o usuário digita a URL (*Uniform Resource Location*) no seu navegador;
- o navegador faz uma consulta ao DNS pelo endereço IP correspondente à URL, se ele não o tiver;
- é iniciada uma conexão TCP através de um aperto de mão de três vias, onde são negociados parâmetros da conexão TCP;
- negociação de TLS, onde cliente e servidor negociam qual será o padrão de criptografia de dados;
- navegador realiza a requisição HTTP;
- servidor responde a requisição;

Referências

- Akhmed, Kamran. How does the internet work?
<https://roadmap.sh/guides/what-is-internet>
- Shuler, Rus. How does the internet work? 2002
<https://web.stanford.edu/class/msande91si/www-spr04/readings/week1/InternetWhitepaper.htm> (necessita atualização)
- Mestrado em Redes e Serviços de Comunicação - Faculdade de Engenharia da Universidade do Porto.
https://paginas.fe.up.pt/mrs01003/TCP_IP.htm
- CloudFlare. What is in an HTTP request?
<https://www.cloudflare.com/en-gb/learning/ddos/glossary/hypertext-transfer-protocol-http/>
- Mozilla Developer Network. An overview of HTTP. 2023.
<https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview>

Referências

- Garsiel, Tali e Irish, Paul. How browsers work. 2011. <https://web.dev/howbrowserswork/>
- Mozilla Developer Network. Populating the page: how browsers work. 2023. https://developer.mozilla.org/en-US/docs/Web/Performance/How_browsers_work