

# **FUNDAMENTOS DE INTERNET E WEB**

**A World Wide Web - WWW**



# ROTEIRO

- **Como surgiu a WWW?**
  - Contexto
- **Os principais componentes da Web**
  - Servidores Web
  - O protocolo HTTP
  - Browsers (Navegadores)/Clientes
  - Servidores de Nomes

# Como surgiu a WWW?

- As ideias por trás da WWW iniciaram-se na década de 1980 no CERN na Suíça
- O desejo de tornar mais fácil as pesquisas por documentos, feitas pelos cientistas, fez em 1989 um pesquisador chamado Tim Berners-Lee iniciar um projeto para gerenciar melhor as informações trocadas entre os pesquisadores.
- Berners-Lee se baseou no sistema chamado ENQUIRE para junto com Robert Cailliau rascunhar a **WWW**

# Como surgiu a WWW

- **1990** - Neste período Berners-Lee usou um computador NeXTcube para escrever o primeiro servidor web e o primeiro navegador chamado de **WorldWideWeb**.
- O NeXTCube era um computador criado pela NeXT, empresa de Steve Jobs, criada em 1985
- **1991** - Nasce a WWW
  - Junto também aparece um novo protocolo (o HTTP), pois não havia protocolos que atendiam às necessidades da comunidade científica.
  - Hipertexto já era discutido desde de 1960, mas o que Lee fez foi unir o hipertexto e a Internet.
  - Lee quis "casar" as duas coisas, mas como ninguém implementou sua ideia, ele decidiu implementar o projeto sozinho.
  - Neste mesmo período ele também desenvolveu um sistema de identificação global e único de recursos, o (URI) Uniform Resource Identifier.

## Como surgiu a WWW

- Em 1991, já havia outros protocolos, como os de serviços de e-mail.
  - Tais protocolos possuíam limitações
- Para suprir algumas limitações como o acesso simples e por links, o protocolo necessitado para acesso à informação deveria atender aos seguintes requisitos:
  - Incorporar a funcionalidade do FTP (File Transfer Protocol)
  - Ser hábil para realizar uma busca
  - Que houvesse um formato de dados automático para negociar a troca de informações
  - Que o cliente/usuário pudesse referenciar outro servidor.

# Funcionamento Básico da Web

- Todos os dados que **"consumimos"** na WWW estão armazenados em servidores.
- Eles são chamados de Servidores Web.
  - Exemplo: **Apache**
- O servidores armazenam os arquivos (de vários tamanhos, tipos, extensões) que são requisitados pelos usuários
- **Importante:** O **usuário/cliente/browser** não consegue acessar diretamente o arquivo que deseja ler sem pedir para o servidor.
  - Aqui o servidor é tanto o equipamento (máquina com disco, processador e memória, interface de rede) e também um programa que fica **"escutando"** o tempo todo, pedidos de quem deseja acessar o conteúdo que está disponível para ler/escrever

# Funcionamento Básico da Web

- É necessário:

- Recursos (áudio, vídeo, texto), os quais são obtidos por meio de uma URL e armazenados em servidores.
- Ao digitar uma URL no navegador desejamos acessar um recurso endereçável.
  - Parte da URL referente ao servidor de rede é separada e transformada em um endereço IP, por um banco de dados da Internet chamado **Domain Name System** (DNS).
  - O navegador estabelece, então, uma conexão TCP com o servidor web localizado no endereço IP retornado.
  - O navegador envia uma requisição HTTP ao servidor para obter o recurso indicado pela parte restante da URL (menos a parte do servidor).
  - O recurso é recebido e interpretado pelo navegador, que realiza então requisições adicionais para figuras, arquivos de formatação, arquivos de script e outros recursos que fazem parte da página.
  - O navegador, então, reconstitui a página na tela do usuário
- **A seguir, vamos discutir de forma separada sobre os servidores Web, HTTP e o DNS, pois são esses os componentes primordiais da Web**

# Servidores Web

- Importante lembrar que a Web baseia-se fundamentalmente em três padrões:
  - **URL** - um sistema que especifica como cada página de informação recebe um "endereço" único onde pode ser encontrada.
  - **HTTP** - um protocolo que especifica como o navegador e servidor web se comunicam entre si.
  - **HTML** - uma linguagem de marcação para codificar a informação de modo que possa ser exibida em uma grande quantidade de dispositivos.



# Servidores Web

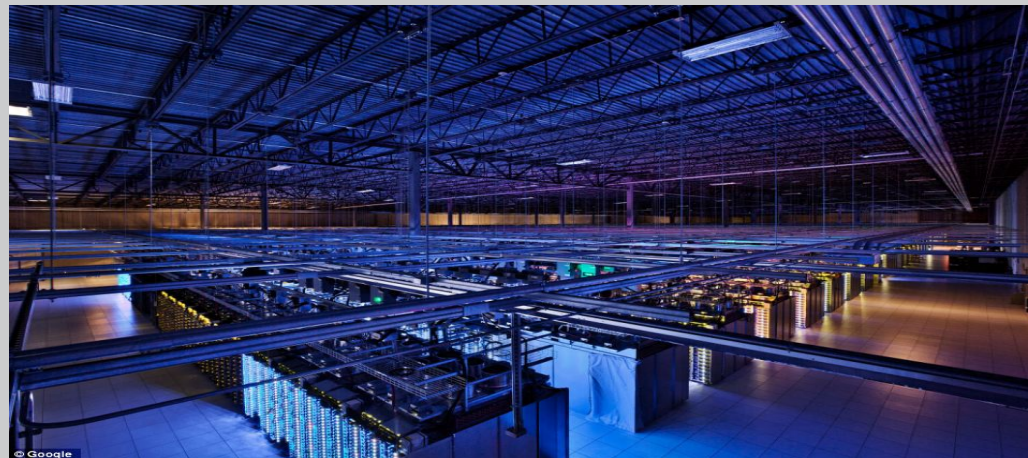
- Um servidor Web é basicamente **um programa** que escuta os pedidos dos navegadores ou aplicações clientes e realiza a execução desses pedidos
- Tais pedidos são páginas web, tais como documentos em HTML com objetos embutidos (imagens, etc.) no servidor
  - O Apache é um servidor web de plataforma aberta, sendo um dos mais populares atualmente. É ativamente mantido pelo Apache Software Foundation.
  - O Apache possui módulos que adicionam mais funções além de apenas servir o conteúdo, mas como fazer isso com segurança, por exemplo. O ***mod\_ssl*** é um utilizado para ativar o suporte ***SSL (Security Socket Layer)***
  - Também possui recursos para: ***balanceamento de carga, reescrita de URL, geolocalização baseada em IP, etc.***

# Servidores Web

- Primeiro Servidor Web



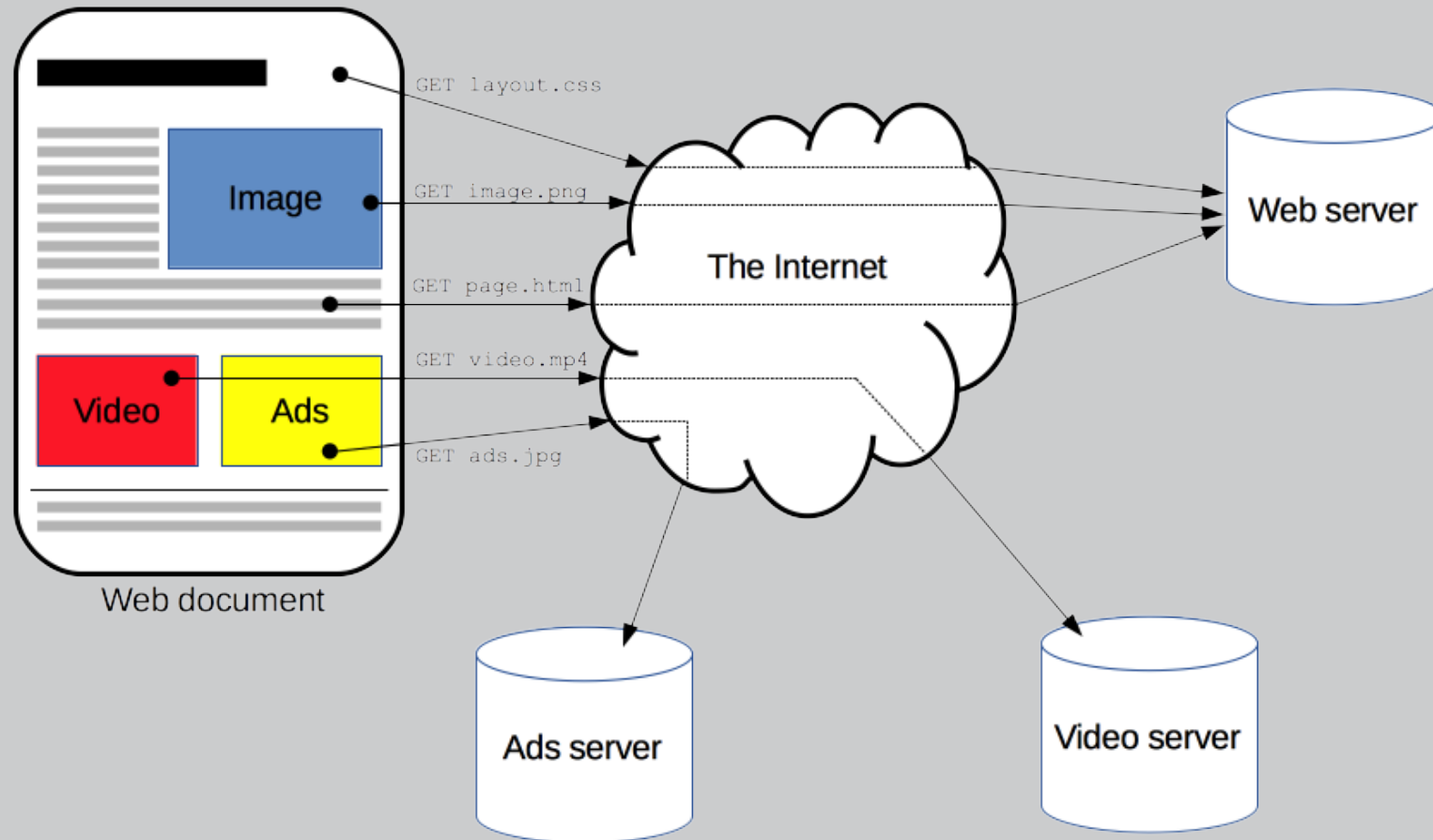
- Data Center do Google com milhares de servidores



# O protocolo HTTP

- Para acessar os recursos disponíveis nos servidores Web, é preciso da ajuda de um protocolo.
  - O protocolo padrão da WWW é o **HTTP (Hypertext Transfer Protocol)**
  - O HTTP é um protocolo cliente-servidor que permite a obtenção de recursos, tais como documentos HTML
  - É um protocolo extensível da camada de aplicação do modelo TCP/IP que é encapsulado sobre o TCP (protocolo confiável como já vimos)
  - É usado para não apenas buscar documentos de hipertexto, mas também imagens e vídeos ou publicar conteúdo em servidores, como nos resultados de formulário HTML (*vamos estudar os elementos `<html>` e `<form>` nas próximas semanas*).

# O protocolo HTTP



# O protocolo HTTP

- **Vejamos um pouco da estrutura do HTTP e o seu funcionamento**
  - **Componentes**
    - **Fundamentalmente o cliente (agente-usuário)**
    - **O servidor de páginas Web**
  - **As três etapas básicas de uma comunicação HTTP são:**
    1. **Estabelecimento da conexão TCP**
    2. **Envio do pedido (requisição) da parte do cliente**
    3. **Envio dos dados da parte do servidor.**

# O protocolo HTTP

- Quem controla uma conexão para encaminhar mensagens HTTP é a camada transporte, ou seja independe do HTTP.
- O HTTP apenas requer que essa conexão seja confiável (sem pelo menos apresentar erros).
  - É por isso que a base de transporte do HTTP é o TCP e não o UDP
  - A versões atuais do HTTP baseiam-se em trazer numa única conexão TCP, praticamente todo o conteúdo de uma página (especialmente se ela possui poucos recursos e se eles estão no mesmo servidor Web).
    - **VANTAGEM: Isso economiza a abertura de novas conexões TCP.**
- Lembre-se, garantir a confiabilidade de entrega de segmentos tem um custo computacional e também de conectividade. Por essa razão o TCP é mais lento que o UDP

# O protocolo HTTP

- Fluxo Padrão

1. Abertura de conexão TCP

2. Envio de mensagem HTTP

*GET / HTTP/1.1*

*Host: developer.mozilla.org*

*Accept-Language: en*

3. Retorno da resposta do servidor

*HTTP/1.1 200 OK*

*Date: Sat, 28 Aug 2020 14:28:02 GMT*

*Server: Apache*

*Last-Modified: Tue, 01 Dec 2009 20:18:22 GMT*

*ETag: "51142bc1-7449-479b075b2891b"*

*Accept-Ranges: bytes*

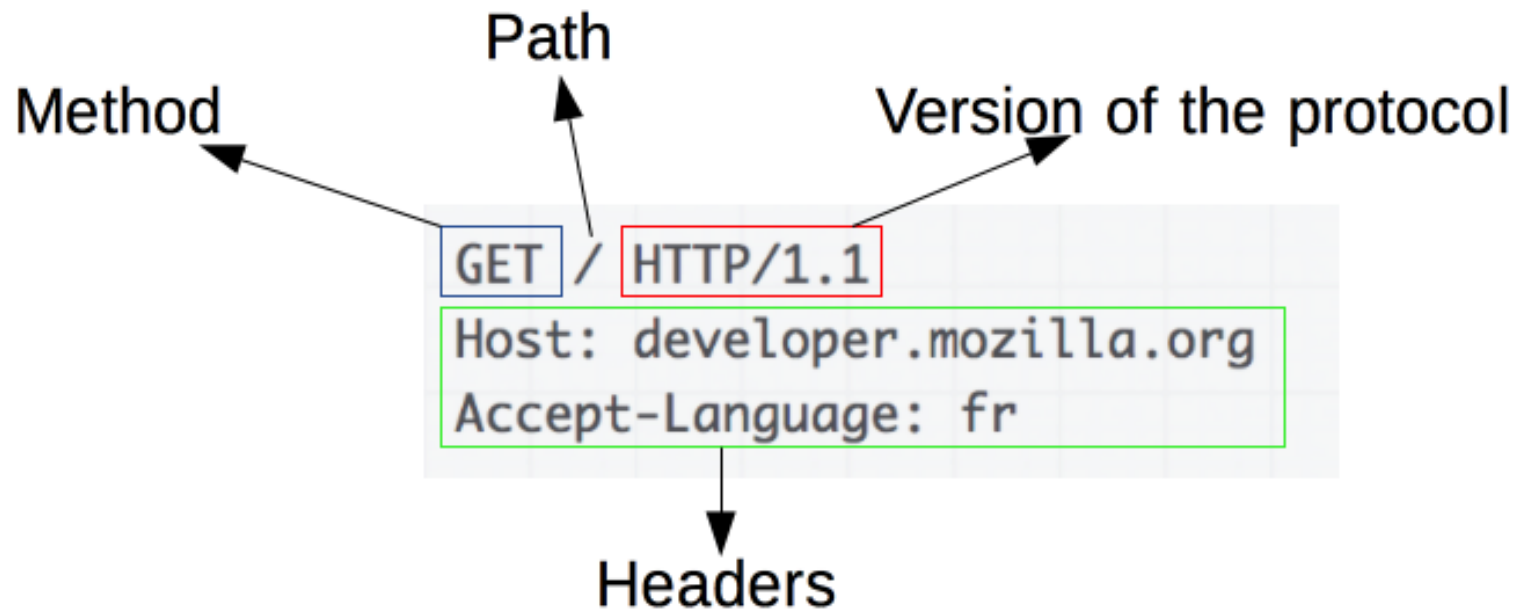
*Content-Length: 29769*

*Content-Type: text/html<!DOCTYPE Texto em html...*

4. Fechamento da conexão TCP

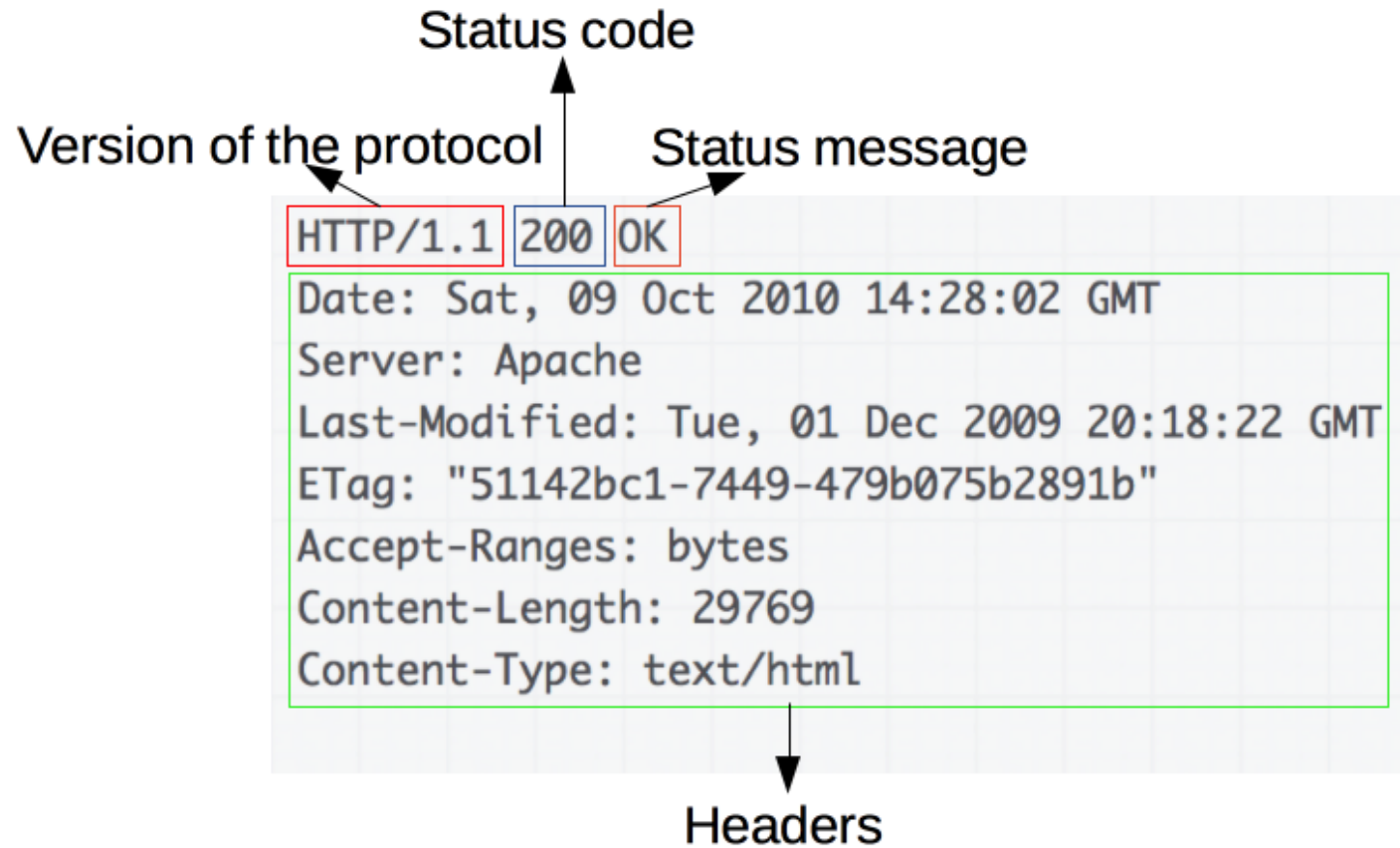


# Requisição HTTP





# Resposta HTTP



# O DNS

- O Domain Name System é uma sigla para sistema de nomes de domínio.
- É um registro que contém nomes de sites e seus respectivos endereços IP associados.
- Essa correlação favorece a transferência de dados entre computadores e permite o acesso à WWW.
- Na barra de endereços de um navegador, o domínio é o nome do site (**univesp.br**), e o servidor de nomes armazena um conjunto deles.

# O DNS

- **Porque o DNS é tão importante e necessário para a WWW?**
  - Os Hosts (computadores, tablets, celulares, etc.) na Internet se comunicam por números (endereços IP).
  - Os sites da web são armazenados em máquinas especiais, os chamados servidores, por meio de serviços de hospedagem.
  - Como já sabemos, cada servidor deve ter um único endereço que permita que as pessoas encontrem esse site e consigam acessar seus dados.
  - Os endereços web contêm uma combinação de IP e um domínio associado.
    - Imagine termos que lembrar os endereços IPs dos sites que visitamos, das lojas on-line que compramos produtos?
    - Um sistema de mapeamento de endereços para nomes ajuda os seres humanos a lembrar mais facilmente no momento da "navegação" para encontrar os recursos (fotos, vídeos, textos, produtos) que deseja.
  - Além disso, um sistema de links baseado em nomes facilita a construção de aplicações que "rodam" na Web. O nome que informamos na barra de endereços do browser é chamado de domínio, como por exemplo: [univesp.br](http://univesp.br)

# O DNS

- Para saber mais sobre DNS sugiro a leitura do site:
  - *[https://www.verisign.com/pt\\_BR/website-presence/online/how-dns-works/index.xhtml](https://www.verisign.com/pt_BR/website-presence/online/how-dns-works/index.xhtml)*

# Referências

- <https://developer.mozilla.org/pt-BR/docs/Web/HTTP/Overview>
- [https://www.gta.ufrj.br/grad/03\\_1/http/](https://www.gta.ufrj.br/grad/03_1/http/)
- [https://pt.wikipedia.org/wiki/World\\_Wide\\_Web#cite\\_note-13](https://pt.wikipedia.org/wiki/World_Wide_Web#cite_note-13)
- [https://pt.wikipedia.org/wiki/Servidor\\_web](https://pt.wikipedia.org/wiki/Servidor_web)
- <https://aws.amazon.com/pt/route53/what-is-dns/>
- [https://www.verisign.com/pt\\_BR/website-presence/online/how-dns-works/index.xhtml](https://www.verisign.com/pt_BR/website-presence/online/how-dns-works/index.xhtml)
- <https://www.datacenterknowledge.com/archives/2011/08/01/report-google-uses-about-900000-servers>

# **FUNDAMENTOS DE INTERNET E WEB**

**A World Wide Web - WWW**

