

ALGORITMOS E PROGRAMAÇÃO DE COMPUTADORES II

Conceitos fundamentais da Web

WWW e Internet

A Internet é uma rede global que conecta computadores ao redor do globo. Permite a troca de mensagens entre programas conectados a ela.

A World Wide Web (WWW ou simplesmente Web) é um sistema distribuído de documentos ligados por hiperlinks e armazenados em servidores Web espalhados pela Internet.

Servidor e cliente

A comunicação entre programas tipicamente ocorre quando um dos programas requisita um recurso a outro programa.

O provedor do recurso é chamado de servidor Web e quem requisita/recebe o recurso é chamado de cliente.

O navegador é um cliente Web, e além de requisitar um recurso, também processa o recurso, exibindo o conteúdo na tela.

URL

Para fazer a requisição de um recurso, é necessário:

- Identificar o recurso**
- Localizar o recurso**
- Usar uma linguagem padronizada de comunicação**

A Web oferece funcionalidades para identificar e localizar um recurso, além de um protocolo de comunicação que permite a troca de mensagens entre cliente e servidor.

URL

A URL (Uniform Resource Locator) oferece um identificador único para cada recurso na Web, definindo também sua localização e o protocolo de comunicação.

Exemplo:



The diagram shows the URL `http://www.w3.org/Consortium/mission.html` with brackets underneath identifying its parts: `http` is the **scheme**, `www.w3.org` is the **host**, and `/Consortium/mission.html` is the **path**.

Fonte: Perkovic, 2015

Scheme

O esquema define como acessar o recurso. No exemplo anterior, usamos o protocolo HTTP.

Existem outros protocolos conhecidos:

- HTTPS
- FTP
- MAILTO
- FILE
- etc.

Host

O host é o nome do servidor que armazena o recurso/documento.

O host é transformado em um endereço IP a partir de uma consulta a um servidor de nomes (DNS), realizada pelo cliente.

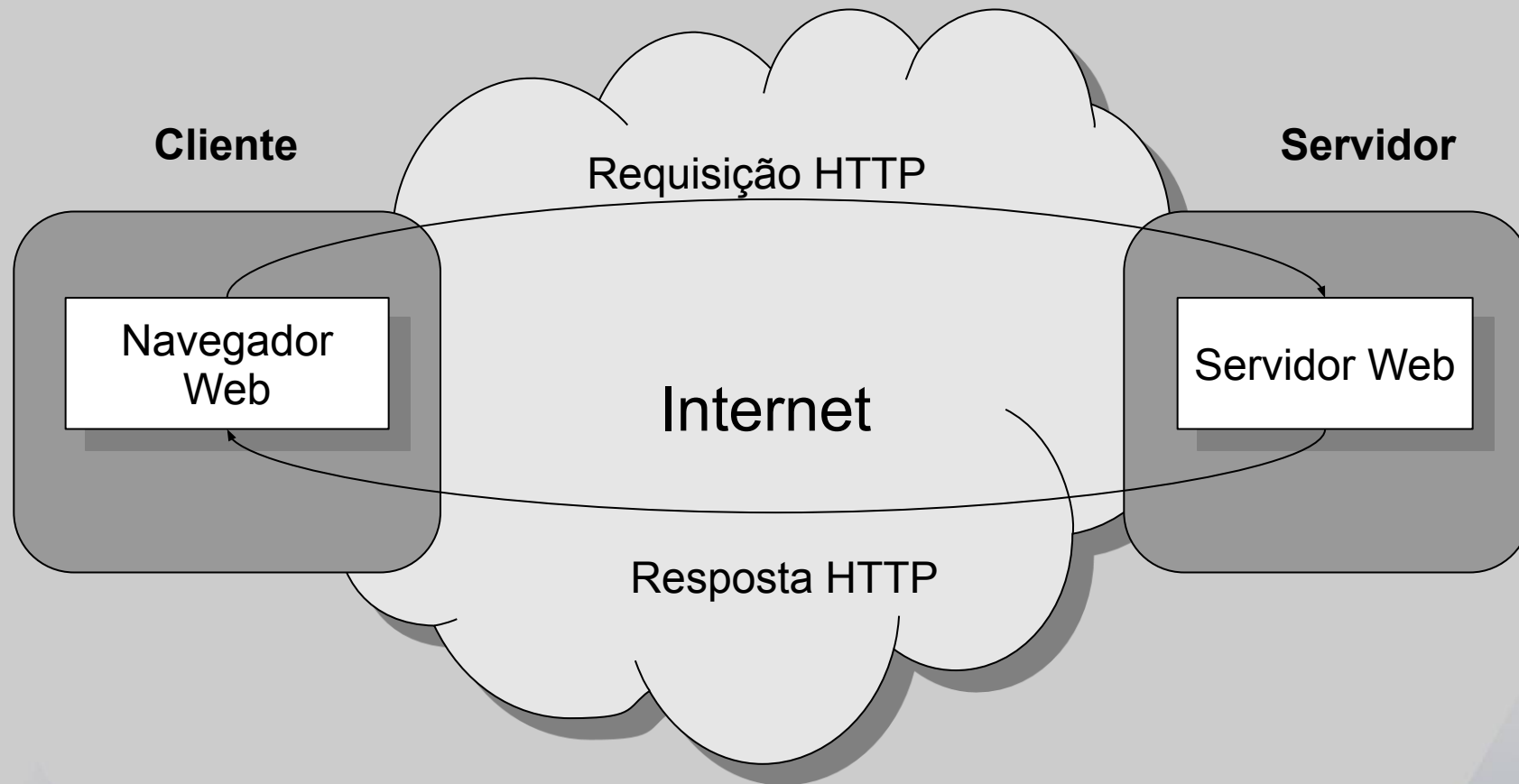
Tanto o hostname como o IP são únicos e localizam o servidor na rede.

Path

O path em uma URL especifica o caminho relativo do recurso a partir de um diretório predefinido (diretório raiz do servidor Web).

Uma vez localizado o servidor, é necessário especificar em qual subdiretório do servidor o recurso se encontra.

HTTP



HTTP

Padrão de mensagens requisição/resposta entre cliente e servidor

Por padrão, usam-se conexões TCP/IP na porta 80

Mensagem típica de requisição:

```
POST / HTTP/1.1
Host: www.somesite.com
Connection: close
Accept: text/plain: text/html
Accept-Language: em-us,en
Accept-Charset: ISO-8859-1,utf-8
User-Agent: Mozilla/5.0 Gecko/20041107 Firefox/1.0
Referer: http://web-sniffer.net
Content-Length: 16
Content-Type: application/x-www-form-urlencoded

productId=ABC123
```

HTTP

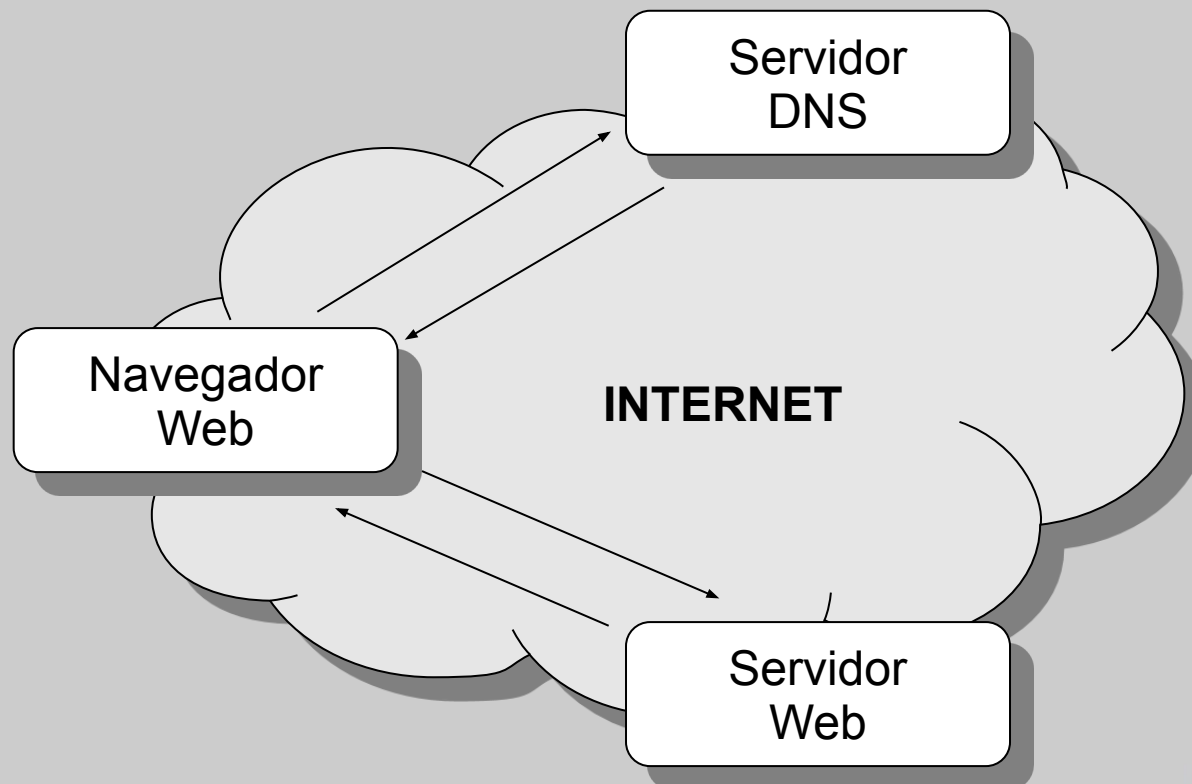
Mensagem típica de resposta:

```
HTTP/1.1 200 OK
Date: Sun, 13 Mar 2005 22:07:43 GMT
Server: Apache/2.0.49
Last-Modified: Sun, 17 Oct 2004 00:26:16 GMT
Content-Length: 70
Keep-Alive: timeout=15, max=99
Connection: Keep-Alive
Content-Type: text/html;charset=UTF-8
```

```
<html>
<head>
<title>
...
```

Fluxo de requisições

Passos para requisição de uma página Web



ALGORITMOS E PROGRAMAÇÃO DE COMPUTADORES II

Conceitos fundamentais da Web