

Arquitetura da World Wide Web

World Wide Web

- Sistema de informação em escala global acessível em tempo real através de redes de computadores como a Internet.

WWW: Usos

- Biblioteca:
 - meio de publicação e compartilhamento de informações instantâneo de amplo alcance e baixo custo.
- Canal comercial:
 - meio para divulgação de produtos e realização de transações comerciais de baixo custo, atingindo um público eclético e de alto poder de compra.

WWW: Histórico

- Bush, 1945
 - introduziu o conceito de hipertexto:
 - modelo que organiza o conhecimento como um grafo onde os vértices são documentos e as arestas são referências cruzadas entre documentos relacionados.
- Nelson, 1965
 - introduziu os termos hipertexto e hipermedia.

WWW: Histórico

- 1970-1990:
 - Uso do hipertexto limitado por dependência de aplicações e gerenciamento uniforme de documentos por diferentes aplicações.
- 1990:
 - WWW amplia capacidade de endereçamento dos hipertextos para a Internet.
- 1994:
 - Criado o World Wide Web Consortium (W3C)

WWW Tecnologias Fundamentais

- Endereçamento
 - *Como encontrar documentos*
 - URI/URL
- Protocolo
 - *Como obter documentos*
 - HTTP
- Formatação
 - Como apresentar documentos
 - HTML

WWW: Endereçamento

- URI (*Universal Resource Identifier*)
 - Cadeia de caracteres que identifica recurso físico ou virtual na WWW
 - Universais:
 - identificam endereços independentemente da rede, protocolo ou servidor
 - Determinísticos:
 - cada objeto na WWW é identificado unicamente por um URI

WWW: Endereçamento

- Protocolos:
 - NNTP: news://domain.com/news.group
 - HTTP: http://www.domain.com/
 - FTP: ftp://ftp.domain.com/pub/
 - SMTP: mailto:user@host
 - telnet: telnet://host.domain/

WWW: Endereçamento

- Classificação:
 - Nomes:
 - URN (*Universal Resource Names*)
 - identificam um recurso independente da sua localização
 - Localizadores:
 - URL (*Universal Resource Locators*)
 - identificam um recurso em função da sua localização

Protocolo

- HTTP
 - nível aplicação
 - suporta formatos diversos:
 - texto
 - imagens
 - áudio
 - não possui noção de estado

HTTP

- Funcionamento:
 - requisição-resposta
 - HTTP 1.0 X HTTP 1.1
- Formato:
 - cabeçalho
 - corpo

HTTP

- Métodos
 - GET
 - HEAD
 - POST
 - PUT
 - DELETE

Formatação

- HTML
 - baseado no SGML
 - elementos de formatação
 - cabeçalhos
 - parágrafos
 - links
 - listas

Arquitetura da WWW

- Clientes
- Servidores
- Meio de comunicação
- Informações

Clientes

- Permite o acesso a informações armazenadas na WWW
- Páginas: conjunto de arquivos em vários formatos
- Tarefas:
 - resolução de nome
 - apresentação

Servidores

- Responde a requisições WWW, enviando páginas dinâmicas ou estáticas
- Processam requisições em paralelo
 - pool de processos
 - criação de processos sob demanda
 - eventos

Meio de comunicação

- Internet: baseada em TCP/IP
- TCP:
 - enfatiza correção e tolerância a falhas
 - ambiente WWW
 - muitas conexões curtas
 - fontes de latência
 - protocolo de estabelecimento de conexão
 - controle de congestionamento

Evolução da WWW

- Escalabilidade
- Conteúdo dinâmico
- Estruturação do conteúdo
- Segurança

Escalabilidade

- Cache:
 - armazena conteúdo de respostas
 - serve respostas armazenadas
 - gerencia espaço de armazenamento
 - localização
 - cliente
 - servidor
 - rede

Caches WWW

- Vantagens
 - redução na latência dos clientes
 - redução de carga nos servidores
 - maior disponibilidade dos documentos
 - diminuição do tráfego na rede
- Desvantagens
 - respostas desatualizadas
 - controle de acesso

Conteúdo Dinâmico

- Documentos cuja geração é regida por parâmetros enviados na própria requisição.
- Estratégias:
 - CGI (Common Gateway Interface)
 - servidores especializados
 - API dos servidores

Estruturação de Conteúdo

- XML: meta-linguagem para definição de linguagens de marcação de documentos estruturados.
- Motivação:
 - permitir o acesso a dados altamente estruturados a partir da WWW.

XML vs. HTML

- Marcadores não são fixos
- Recursos de definição de marcadores e relações estruturais
- Recursos para definição da semântica de marcadores e da sua forma de apresentação (estilo)

Segurança

- SSL (Secure Socket Layer)
 - permite interação segura entre o cliente e o servidor, deixando explícito apenas a informação para encaminhamento da requisição.
 - conceitos:
 - sessão
 - conexão

SSL

- Sessão
 - associação entre cliente e servidor
 - define parâmetros de segurança
 - autenticação mútua
 - negociação das chaves de criptografia
 - evita negociação a cada conexão
- Conexão
 - utiliza uma chave assimétrica negociada