

# Základy počítačových sítí

## Webové služby



Základy počítačových sítí  
Lekce 7  
Ing. Jiří ledvina, CSc

## Úvod



- Historie
- Webové komponenty
- HyperText Markup Language (HTML)
- Uniform Resource Locator (URL)
- Postup vytváření HTML dokumentu
- Statické, dynamické a aktivní stránky
- Hypertext Transport Protocol (HTTP)
- Cookies, vyrovnávací paměti, proxy,
- Vyhledávání a indexování
- RSS
- Bezpečný přenos dat, HTTPS

## Historie WWW



- Vytvořen Tim Berners-Lee v letech 1989 až 1990 v CERN (Evropská laboratoř pro fyziku částic)
- 1994 – Mark Andreesen vymyslel v NCSA (National Center for Super Computing Applications) MOSAIC
  - první grafický prohlížeč
  - první Internetová "killer application" – první opravdová aplikace, pro kterou začal opravdu Internet používat
  - volně přístupná
  - později Netscape Inc.

29.11.2006

Základy počítačových sítí - lekce 7

3

## Historie WWW



- 1995 – webové přenosy se stávají dominantními
  - exponenciální nárůst provozu na síti
  - elektronická komerce (E-commerce)
  - WWW konsorcium
- Tim Berners-Lee
  - Fyzik, ne počítačový specialista
  - Sdílení dat z fyzikálních experimentů
  - Protože FTP bylo příliš obtížné
  - Prostředek pro přenos textu i grafiky najednou
  - Nyní strategie "ukaz a klikni"

29.11.2006

Základy počítačových sítí - lekce 7

4



## Webové komponenty

- Prohlížeč
  - Webový klient. Nyní se upouští od označení prohlížeč, protože webový dokument se může „zobrazit“ i zvukově, ...
  - Internet Explorer, Firefox, Mozilla, Netscape, Opera, Mosaic, lynx
- Webový server
  - projekt Apache, Jakarta, Microsoft
- Reprezentace dokumentů (HTML)
  - Text, obraz, zvuk, video
- Identifikace dokumentů (URL)
- Přenosový protokol (HTTP)
  - K přenosu se využívá spolehlivý protokol TCP

29.11.2006

Základy počítačových sítí - lekce 7

5



## Webový klient (browser)

- aplikační program
- představuje uživatelské rozhraní pro přístup k Webu
- stahuje informace z webového serveru
- zobrazuje stažené informace

29.11.2006

Základy počítačových sítí - lekce 7

6



## Webový server

- úložiště webových dokumentů
- odpovídá na požadavky prohlížeče a posílá mu kopie dokumentů
- Spolupracuje s jinými servery při dynamickém generování dokumentů (jízdni řád, elektronické obchodování, STAG, ...)



## Webový dokument

- webový dokument
  - Text, obrázky, zvuk, video
  - odkazy na ostatní webové stránky
- webový dokument a hypertextové odkazy
  - Hypertextový odkaz je spojen s objektem nebo oblastí na obrazovce
  - Vnitřně se jeví jako symbolický link
  - Výhoda - může odkazovat na dokument na jiném počítači
  - Nevýhoda - nemusí platit (neplatné URL)

## Webový dokument



- je označován jako webová stránka
- jednu webovou stránku tvoří jeden soubor
  - Používá se přípona \*.htm, \*.html
- může obsahovat
  - textový soubor
  - binární obrázek
- text je standardizován
  - známý jako HTML (HyperText Markup Language)
  - obsahuje ASCII znaky nebo znaky národních abeced
- HTML specifikuje obsah a rozvržení textu
  - Způsob zobrazení závisí na webovém klientovi

29.11.2006

Základy počítačových sítí - lekce 7

9

## HTML (HyperText Markup Language)



- Vychází z obecného jazyka pro popis dokumentů SGML (Standard Generalized Markup language)
- Jazyk pro popis obsahu a rozvržení dokumentu
- Na způsobu napsání dokumentu nezáleží (mezery a nové řádky neovlivní zobrazení – mohou se použít pro zvýšení přehlednosti zápisu dokumentu)
- Způsob zobrazení je dán zabudovanými značkami (tag)
- Značky jsou párové nebo nepárové
  - Formát značky
    - počáteční `<TAGNAME>`
    - koncová `</TAGNAME>`
  - Příklad – **tučný text kurzívou**
    - `<B><I> tučný text kurzívou</I></B>`

29.11.2006

Základy počítačových sítí - lekce 7

10

## Obecný formát HTML dokumentu



```
<HTML>
  <HEAD>
    <TITLE>
      text který se zobrazí jako titulek dokumentu
    </TITLE>
      Další informace v záhlaví
  </HEAD>
  <BODY>
      tělo dokumentu, jeho obsah se zobrazí
      jako webová stránka
  </BODY>
</HTML>
```

29.11.2006

Základy počítačových sítí - lekce 7

11

## Typický příklad záhlaví



- Generováno editorem HTML stránek FrontPage 5.0

```
<!DOCTYPE html PUBLIC "-//w3c//dtd html 4.0 transitional//en">
<html> <head>
<meta http-equiv="Content-Language" content="cs">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1250">
<meta name="GENERATOR" content="Microsoft FrontPage 5.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<meta name="Author" content="Carl Ellison">
<meta name="Keywords" content="X.509, PGP, SPKI, SDSI">
<meta name="Microsoft Theme" content="waves 011">
<title>Porovnání certifikátů</title>
</head>
..... Tělo dokumentu .....
</html>
```

- Výrazně části záhlaví udávají použitý jazyk a znakovou sadu

29.11.2006

Základy počítačových sítí - lekce 7

12



## Příklady HTML značek

- Začátek odstavce (nepárová)  
`<P>`
- Přejít na novou řádku (nepárová)  
`<BR>`
- Nadpis 1 (největší písmo)  
`<H1> ...text... </H1>`
- Nadpis 2 (menší)  
`<H2> ...text... </H2>`
- komentář  
`<!-- ... /-->`
- Tučné písmo  
`<B> ...text... </B>`
- Kurzíva  
`<I> ...text... </I>`
- Podtržené písmo  
`<U> ...text... </U>`
- Seznam (jeden prvek)  
`<ul> <li> </li> </ul>`
- Číslované seznamy  
`<ol> <li> </li> </ol>`



## Linky (odkazy)

- používají se značky `<a>` a `</a>`
- **relativní linky**
  - odkazují na stránku vztahenou k tomuto dokumentu
  - používají se pro zachování přenositelnosti dokumentů
  - např. (zvýrazněné se zobrazí v dokumentu, podtržené je odkaz)  
`Výsledky zkoušky ze dne <a href="/vysledky/18.11.2005.html">18.11.2005 </a>`
- **absolutní linky**
  - odkazují na cizí dokumenty
  - používají se pro přístup k dokumentům na „cizích“ serverech
  - např. (zvýrazněné se zobrazí v dokumentu, podtržené je odkaz)  
`Výsledky zkoušky ze dne <a href="http://home.zcu.cz/~novak/vysledky/18.11.2005.html">18.11.2005 </a>`



## Ukotvení (anchor)

- Zakotvení (anchor) – přechod na určené místo v dokumentu
  - může být umístěno kdekoli v dokumentu  
`<a name="navez_znacky"> Pozice značky</a>`
  - přechod na značku v tomtéž dokumentu  
`<a href="navez_znacky"> Přechod na značku</a>`
  - přechod na značku z jiného dokumentu  
`<a href="cesta k dokumentu#navez_znacky"> Přechod na značku</a>`
  - cesta k dokumentu může být relativní nebo absolutní

29.11.2006

Základy počítačových sítí - lekce 7

15



## Vkládání obrázků

- explicitně označeno jako obrázek
- specifikace pomocí `<IMG SRC="jmeno_souboru">`
- lze specifikovat i další parametry, např. zarovnání  
`<IMG SRC="jmeno_souboru" align=middle>`
- Nebo rozměr obrázku a náhradní text pokud klient neumí obrázek zobrazit  
``

29.11.2006

Základy počítačových sítí - lekce 7

16





## Kaskádové styly - CSS

- Od popisu stránek přímo pomocí HTML značek se upouští pro malou pružnost při provádění dodatečných úprav
- Zavádí se kaskádové styly (Cascading Style Sheets - css)
- Nyní již ve verzi 3
- Používá značku <style>

```
<style>  
  Selektor {vlastnost:hodnota; vlastnost:hodnota}  
  Selektor {vlastnost:hodnota}  
</style>
```



## Kaskádové styly - CSS

- Příklady zápisu
  - přímé v dokumentu (style = )  
`<p style="text-align: center">Text odstavce ... .. </p>`
  - v hlavičce dokumentu <style> ... </style>

```
<head>  
<title> ... </title>  
<style type="text/css">  
h2 {color: blue; font-style: italic}  
</style>  
</head>  
<body>
```

```
<h2>Nadpis </h2>  
<body>
```

## Kaskádové styly - CSS



- Příklad zápisu

- v externím souboru \*.css

```
<link rel="stylesheet" href="soubor.css">
```

- nebo

```
<style>
```

```
@import url("soubor.css")
```

```
</style>
```

## Kaskádové styly - CSS



- Možnosti CSS (některé)

- Jednotná změna fontu
- Jednotný formát odstavce
- Jednotná manipulace s barvami
- Nastavení velikosti a obtékání
- Nastavení okrajů
- Jednotné seznamy
- Jednotné tabulky

- Výhoda spočívá v tom, že určíme atribut, kterému přiřadíme definici vlastností. Pokud chceme vlastnosti změnit, stačí tak učinit na jednom místě

- Více na <http://www.jakpsatweb.cz/>

## XHTML



- Nová norma HTML
- Vývoj HTML skončil verzí 4.01
- X – extensible (rozšiřitelný)
- Zúžení možností HTML z důvodu lepší ověřitelnosti souladu s normou
- Nyní se používá XHTML 1.0 a 1.1
- Určení použitého XHTML (přípustnost tagů a jejich atributů) se definuje na začátku dokumentu např.  

```
<?xml version="1.0" encoding="iso-8859-2"?>  
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```
- Do dokumentu dosadí většinou HTML editor

## Rozdíly HTML a XHTML



- XHTML striktně vyžaduje
  - Všechny atributy mají hodnoty v uvozovkách
  - Zákaz křížení tagů
- Rozdíly mezi HTML a XHTML
  - Tagy a atributy jsou malými písmeny
  - Nepárové tagy končí lomítkem
  - Párové tagy jsou párové povinně
  - Všechny atributy musejí mít hodnotu
  - Interní javascript a styly se zapisují jiným způsobem
  - Dokument má mít XML prolog.
  - Dokument požaduje správný doctype.
- Více na <http://www.jakpsatweb.cz/html/xhtmll.html>

## URL (Uniform Resource Locator)



- Slouží k identifikaci objektu
- Má textovou podobu
- Byl vytvořen pro identifikaci různých objektů, mimo jiné i webových stránek
- Má obecný tvar

protokol://uživatel:heslo@doménové\_jméno:port/cesta\_k\_souboru?parametry

protokol://uživatel:heslo@doménové\_jméno:port/cesta\_k\_souboru#návěští

- Znaky ':', '/', '@', '?', '#' slouží k oddělení a určení jednotlivých částí URL

## URL (Uniform Resource Locator)



- Speciální znaky a jejich význam
  - ':' - oddělení protokolu od jména nebo IP adresy počítače
  - '@' - oddělení uživatelského jména od jména nebo IP adresy počítače
  - '#' - označení odkazu na návěští ve stránce
  - '~' - označení domovského adresáře pro webové stránky uživatele (public\_html)
  - '?' - označení že následují parametry
  - '/' nebo '\' - oddělení jednotlivých podadresářů
  - './' - aktuální adresář
  - '../' - adresář vyšší úrovně (používá se při relativním odkazování)

## URL (Uniform Resource Locator)



- Např. URL  
http:// home.zcu.cz:8080/~novak/soubor.html
- Se chápe následovně:
  - http - protokol
  - 8080 - číslo portu
  - home.zcu.cz - doménové jméno
  - ~novak - cesta k souboru
  - soubor.html - soubor
- Pokud některá část URL chybí, nahradí se předdefinovanou hodnotou
  - protokol - HTTP
  - port - 80
  - soubor - index.htm, index.html, ...

29.11.2006

Základy počítačových sítí - lekce 7

25

## URL (Uniform Resource Locator)



- Protokol určuje způsob přístupu k dokumentu
- Může být (na písmu (velké/malé) nezáleží)
  - HTTP - protokol HTTP
  - HTTPS - zabezpečený HTTP (šifrování)
  - FTP - přístup pomocí FTP
  - FILE - soubor na lokálním disku
  - GOPHER - předchůdce HTTP
  - MAILTO - adresa el. pošty
  - TELNET - vzdálený přístup

29.11.2006

Základy počítačových sítí - lekce 7

26

## Vytváření HTML dokumentu



- jakýmkoliv textovým editorem (Notepad, Wordpad a další)
- Speciálním HTML editory – WYSIWIG (Microsoft FrontPage, Microsoft Office Publisher, DreamWeaver a další – mnohé volně šiřitelné – Nvu „new view“, Mozilla Composer, Netscape Composer, Trellian WebPAGE, )
- Existuje i export stránek z různých WYSIWIG editorů (MS word)
- v počátcích je výhodné používat textový editor nebo jednoduchý HTML editor – pochopení principu, jednodušší konstrukce stránek

29.11.2006

Základy počítačových sítí - lekce 7

27

## Vytváření vlastních webových stránek



- Vytvoření samostatné stránky a její lokální odzkoušení www prohlížečem (file:// cesta k souboru), vytvoření vnitřních odkazů a jejich odzkoušení
- soubor opatřit příponou htm nebo html (dohoda)
- vytvoření dalších stránek, vzájemné propojení stránek relativními odkazy, vytvoření absolutních odkazů na cizí stránky
- lokální odzkoušení vytvořených vazeb

29.11.2006

Základy počítačových sítí - lekce 7

28

## Vytváření vlastních webových stránek



- stránky se ukládají do vhodně vytvořené adresářové struktury, např. html stránky do jednoho adresáře, obrázky do jiného, související dokumenty do dalšího, atd. (obecný předpis neexistuje)
- přesunutí stránek na webový server a odzkoušení webovým prohlížečem (např.  
[http:// počítač. firma.doména/~login\\_name/cesta/ ... /soubor.html](http://počítač.firma.doména/~login_name/cesta/.../soubor.html)  
( ~ znamená značku pro domácí adresáře uživatelů)

## Vytváření vlastních webových stránek



- Domácí adresář pro html stránky je obvykle ~/public\_html
- Pokud není uvedeno jinak (v URL není uveden odkaz na konkrétní dokument), hledá prohlížeč v tomto adresáři soubor index.htm nebo index.html
- např. uživatelské stránky na ZČU:  
[home.zcu.cz/~login\\_name](http://home.zcu.cz/~login_name)
- Prohlížení vytvořených webových stránek
  - zobrazení vybrané stránky prohlížečem
  - volba zobrazit/zdrojový kód



## Typy webových stránek

- **statické**
  - HTML stránky jsou uloženy v souboru
  - jsou neměnné, mohou obsahovat text, obrázky, odkazy, ...
- **dynamické**
  - jsou vytvářeny serverem za běhu, na přání
  - jsou výstupem nějakého programu
  - Např. CGI (Common Gateway Interface)
  - nyní častěji přímé volání programu ze serveru (PHP, Java, ... )
- **aktivní**
  - spuštěny v prohlížeči na straně klienta
  - obsahují program, mohou komunikovat s prostředím (uživatelé)
  - Mohou bezprostředně reagovat na pohyb myši, stisk klávesy
  - používají technologii Java, JavaScript nebo jiné

29.11.2006

Základy počítačových sítí - lekce 7

31



## CGI technologie

- **URL specifikuje**
  - adresu webového serveru
  - CGI program na serveru (název.cgi)
  - argumenty programu (?jméno=hodnota; ... )
- **web server**
  - používá TCP komunikaci
  - přijímá HTTP požadavek od klienta
  - spouští určený CGI program
  - vrací výsledek (textový výstup programu) klientovi

29.11.2006

Základy počítačových sítí - lekce 7

32



## CGI program



- provádí zadaný výpočet
- je často psán ve skriptovacím jazyce
- za běhu produkuje výstupní soubor
- na počátku svého běhu generuje hlavičku
- hlavička obsahuje informace ve tvaru klíčové slovo:informace, např.
  - Content Type: text/html; charset=UTF-8 - HTML dokument
  - Server: GWS/2.1 - informace o serveru
  - Content-Length: 1000 - délka datové části
  - Date: Thu, 23 Nov 2006 10:53:51 GTM

## Dynamické vytváření stránek



- CGI má velkou režii spojenou s vytvořením obslužného procesu a zavedením programového modulu do paměti
- Proto se častěji používá jiná metoda, kdy interpret jazyka programu pro vytvoření stránky je součástí (modul) webového serveru
- Populárním jazykem je PHP
  - Programový modul se umístí jako zakomentovaný text do HTML stránky
  - Webový server stránku před odesláním analyzuje a zadaný program interpretuje – výstupem je statický text (statická část stránky) a dynamicky programem vygenerovaný text
  - Klient získá HTML dokument, o programu nic neví



## Dynamické vytváření stránek

- Dalším prostředkem je Java
  - Součástí webového serveru je interpret jazyka Java
  - Programy upravené pro volání webovým serverem se označují jako servlety
  - Nejsou součástí webové stránky, jsou umístěny v dohodnutém adresáři
  - Jsou jim předávány dohodnutou metodou i parametry
  - Aby nebylo třeba vše generovat programem, existují JSP (Java Server Page) – část statická (statický text) a část dynamická (servlety) – obdoba PHP

29.11.2006

Základy počítačových sítí - lekce 7

35



## Aktivní stránky

- Program se spouští na straně klienta (v prohlížeči)
- Prohlížeč musí obsahovat interpret jazyka (nejčastěji javascript nebo java)
- Výhoda je možnost reagovat okamžitě na události spojené s pohybem myši nebo zadáním z klávesnice
- Použití – hry, výpočty, bankovníctví
- Javascript je zakomentovanou součástí webové stránky – prohlížeč Javascript interpretuje
- Java se přenese jako samostatný modul – applet
  - V prohlížeči se interpretuje (spustí se) a výsledky zobrazuje do přiděleného prostoru (grafické okénko na webové stránce)

29.11.2006

Základy počítačových sítí - lekce 7

36

# HTTP

## HyperText Transfer Protocol



- HTTP je aplikační protokol, pracuje nad standardní síťovou infrastrukturou (TCP/IP)
- Existují 3 verze
  - 0.9 – původní návrh Berners-Lee
  - 1.0 – (RFC 1945) – používá se nejčastěji
  - 1.1 – (RFC 2068) – počet implementací neustále narůstá
- Komunikační protokol typu server/klient (komunikace typu požadavek – odpověď)
- Bezstavový protokol
  - Server otevře spojení, obslouží požadavek, uzavře spojení

29.11.2006

Základy počítačových sítí - lekce 7

37

## HTTP požadavky (request)



- Klient může poslat serveru požadavek typu
  - GET – požadavek na zaslání dokumentu dle URL
  - PUT – uložení dokumentu určeného URL
  - HEAD – obnova informace o dokumentu dle URL
  - OPTIONS – obnova informace o dostupných volitelných parametrech
  - POST – dodání informace na server
  - DELETE – zrušení dokumentu dle URL
  - TRACE – vrácení zprávy s požadavkem z důvodu ladění
  - CONNECT – používají vyrovnávací paměti (cache)

29.11.2006

Základy počítačových sítí - lekce 7

38

## HTTP požadavky (request)



- HTTP požadavek má tvar  
`<požadavek> URL HTTP <verze>`
- Např.  
`GET http://home.zcu.cz/~novak/index.html HTTP/1.1`
- Následuje záhlaví požadavku (parametry)
- Požadavek PUT má i tělo (text posílaný serveru)

## HTTP odpověď (response)



- HTTP server posílá odpověď ve tvaru  
`HTTP/<verze> xyz kód odpovědi slovně`
- následuje záhlaví s parametry
- Vlastní tělo zprávy s požadovaným dokumentem (odpověď na GET nebo POST)
- V záhlaví je např.
  - Date: Friday, 27-Apr-01 13:30:01 GMT
  - Content-length: 3001

## HTTP kódy odpovědí



- Kód odpovědi je ve tvaru
  - XYZ slovní význam
  - Jsou rozděleny do pěti skupin podle významu
    - 1xx – informační – požadavek byl přijat a zpracovává se
    - 2xx – úspěšné volání – akce byla přijata, akceptována a zpracována
    - 3xx – redirekce – je třeba provést další akci (přesměrování)
    - 4xx – chyba klienta – chybná syntaxe, nemůže být provedeno
    - 5xx – chyba serveru – server nemůže požadavek provést (např. přetížení serveru)
  - Další dvě pozice kódu upřesňují jeho význam

29.11.2006

Základy počítačových sítí - lekce 7

41

## Cookies – záznam stavu



- Vztah mezi klientem a serverem nevyžaduje zapamatování stavu komunikace na straně serveru
  - Výhoda je zjednodušení serveru
  - Zvýšení odolnosti proti zahlcení
  - Nevýhoda že si server nepamatuje, které stránky již uživatel navštívil
- Cookies (koláčky) slouží k zapamatování stavu z pohledu uživatele
  - Generuje je server a posílá klientovi
  - Ukládají se do speciální vyrovnávací paměti na disku
  - V případě potřeby je klient pošle serveru

29.11.2006

Základy počítačových sítí - lekce 7

42



## Cookies – záznam stavu

- Cookies obsahují informace, definované serverem, které by klient neměl měnit
  - Set-cookie: textový řetězec (posílá server)
  - Cookie: textový řetězec (posílá klient)
- Cookies obsahují (přibližně, liší se dle RFC specifikace)
  - Jméno domény – kde se mohou uplatnit
  - Cestu ke stránce – určení dokumentu na serveru
  - Obsah – vlastní rozlišovací informaci ve tvaru název=obsah
  - Dobu expirace
  - Bezpečné spojení ano/ne

29.11.2006

Základy počítačových sítí - lekce 7

43



## Perzistentní spojení

- Původně komunikoval HTTP server s klientem pomocí TCP tak, že se vytvořilo spojení pro přenos pouze jednoho dokumentu
- Obsahoval-li dokument obrázky, vytvořilo se pro přenos každého obrázku další spojení
- Z důvodu snížení režie se přenáší během jednoho spojení celá stránka, tj. jak textová část, tak i obrázky

29.11.2006

Základy počítačových sítí - lekce 7

44



## Vyrovnávací paměti

- Slouží k omezení zbytečných přenosů v síti
- Vyrovnávací paměti (cache)
  - Na straně klienta (disk, paměť počítače)
    - Načtené stránky se ukládají do vyrovnávací paměti
    - Při požadavku opakovaného čtení stránky se zkontroluje není-li již načtena
    - Pokud se její obsah mezi tím nezměnil, načte se z vyrovnávací paměti
    - Ke kontrole slouží příkaz HEAD a porovnání s dobou života dokumentu
    - Ukládání do vyrovnávací paměti lze v dokumentu zakázat (např. při přístupu k bankovnímu účtu), příkazem mazat

29.11.2006

Základy počítačových sítí - lekce 7

45



## Vyrovnávací paměti

- Vyrovnávací paměti (cache)
  - Na komponentách počítačové sítě
    - Konfigurovatelné servery (cache servery)
      - Webový klient může mít nakonfigurovanou adresu proxy serveru, přes který je umožněn přístup z firemní sítě do Internetu
      - Většinou se počítač označuje jako proxy.firma.cz nebo cache.firma.cz a používá implicitní port 3128
      - Součástí proxy serveru mohou být i vyrovnávací paměti
      - Webový klient posílá požadavek na proxy server, ten provede kontrolu dostupnosti dokumentu ve své paměti. Buď na požadavek odpoví sám, nebo jej pošle originálnímu serveru
      - Odpovědi originálního serveru na požadavky automaticky ukládá do vyrovnávací paměti pro další použití

29.11.2006

Základy počítačových sítí - lekce 7

46

## Vyrovnávací paměti



- Vyrovnávací paměti (cache)
  - Na komponentách počítačové sítě
    - Transparentní servery (transparentní cache servery)
      - Jsou umístěny v páteřních částech Internetu
      - Nekonfigurují se (uživatel o nich neví – proto transparentní)
      - Směrovače v Internetu přesměrují automaticky HTTP požadavky na počítače s vyrovnávacími paměťmi
      - Proveďte se kontrola dostupnosti dokumentu a buď se poskytne kopie, nebo se požadavek předá originálnímu serveru
      - Existuje i protokol pro výměnu informací mezi cache servery – výměna zachycených souborů z důvodu dalšího zvýšení průchodnosti

## Proxy



- Proxy znamená zástupce a v počítačových sítích se tento pojem vyskytuje v tomto významu poměrně často
- V tomto případě zprostředkovává proxy server spojení firemní sítě (intranetu) a vnější sítě (Internetem)
- Slouží jako součást ochrany vnitřní sítě před napadením zvenku





## Vyhledávání a indexování

- K poskytování informací nestačí HTTP servery, protože informací je moc
- Dochází k budování indexových serverů, které obsahují seznam dokumentů přístupných podle klíčových slov
- Indexy se vytváří
  - Na přání (manuálním zadáním dokument, klíčová slova)
  - Na základě informací uvedených v hlavičce dokumentu (Keywords=)
  - Automaticky pomocí prohlédávacích strojů – robotů, kteří neustále prohlédávají Internet, hledají HTTP servery a čtou všechny dostupné HTML stránky a třídí je podle slov (klíčových slov), získaných z textu

29.11.2006

Základy počítačových sítí - lekce 7

49



## Vyhledávání a indexování

- Pro získání informací (odkazů na webové stránky) slouží „vyhledávače“, které vyhodnotí zadaný výraz a vrátí relevantní odkazy
- Nejznámější vyhledávač Google
- Indexování a vyhledávání může být realizováno i na jednom webovém serveru (firemní weby – vyhledání informací vztahujících se k výrobku)

29.11.2006

Základy počítačových sítí - lekce 7

50

## Oznamování o změnách stránky



- Při sledování většího počtu webových stránek je problém se zjišťováním jejich změn
- RSS (Really Simple Syndication (0.9x) nebo RDF Site Summary (1.0))
- RDF (Resource Description Framework)
- Cílem je přebírat obsah zdrojů na Internetu a v přehledné formě je nabízet uživateli
- Informaci vytváří autor stránky ve formátu XML (eXtensible Markup Language) a v tomto formátu se přenáší Internetem do RSS čtečky
- RSS čtečka periodicky zjišťuje změny na zadané stránce, stahuje RSS dokument a interpretuje jeho obsah jako seznam změněných dokumentů
- Např. Headline Viewer, FeedReader, AmphetaDesk
- Viz <http://interval.cz/clanky/rss-rss/>

29.11.2006

Základy počítačových sítí - lekce 7

51

## Zabezpečení HTTP



- Přenos pomocí HTTP je otevřený – nelze takto přenášet citlivé informace
- Systém byl doplněn o SSL vrstvu (Secure Socket Layer), která leží mezi TCP a HTTP
- SSL zajišťuje šifrování přenášených dat
  - Je založeno na certifikátech
  - Dovoluje ověřit server (anonymní přístup klienta)
  - Vzájemné ověření serveru i klienta
  - Při ověřování (asymetrická šifra) se přenesou relační klíče (symetrická šifra) pro další komunikaci
- Takto zabezpečený protokol je označován jako HTTPS

29.11.2006

Základy počítačových sítí - lekce 7

52