



9

Servicii web

04 decembrie 2009

Spending an evening on the World Wide Web is much like sitting down to a dinner of Cheetos... two hours later your fingers are yellow and you're no longer hungry, but you haven't been nourished.

Clifford Stoll

- World Wide Web
- Apache Web Server
- Virtual Hosting în Apache

- World Wide Web (Web, W3)
- Serviciu din Internet
- Rețea (web) de informație
 - folosire de hyperlink-uri
- Sir Tim Berners-Lee, 1990, CERN
- W3C (World Wide Web Consortium)
- Probabil cel mai utilizat protocol din Internet
- Searchable Web vs. Invisible Web/Deep Web
- Mai 2009, 109.5 milioane de site-uri

- Pagini web
 - Denumirea resurselor web
- URL
 - Uniform Resource Locator
 - Localizarea unei resurse
 - `protocol://username:password@hostname:port/path/to/resource`
- HTML
 - Hypertext Markup Language
 - Limbajul de descriere a paginilor web (conținut, format, structură)
- HTTP
 - Hypertext Transfer Protocol
 - Protocolul folosit de WWW
 - Sistem client/server
 - Portul 80
 - Închiderea conexiunii

- Utilizatorul deschide clientul (browser-ul, navigatorul)
- Se introduce URL-ul (`http://hostname/path/to/resource`)
- Cerere DNS pentru hostname
- Conectare la serverul web de pe hostname, portul 80
- Se trimite mesajul
 - `GET /path/to/resource HTTP/1.1\r\n`
 - `Host: hostname\r\n\r\n`
- Serverul web citește resursa din `$document_root/path/to/resource`
- Serverul web răspunde cu un antet (200 OK, 300 Redirect, 404 No such resource etc.) și resursa
- Clientul/browserul redă pagina (HTML rendering)
- Clientul/browserul închide conexiunea

- Cunosc protocolul HTTP (HTTPS)
- Implicit, ascultă conexiuni pe portul 80
- Apache (cotă de piață, funcționalitate)
- lighttpd (eficiență)
- nginx (eficiență)
- Microsoft IIS
- Containere: Apache Tomcat

- Browsere, navigatoare
- Dezvoltate intens; utilizate des
- Rendering engine
- ACID tests
- Internet Explorer, Firefox, Safari, Opera, Chrome
- CLI: lynx, w3m, links, elinks
- Neinteractive: wget, curl (libcurl)

- CGI, FGCI
- Module pentru servere
- Javascript, AJAX
- HTTPS, SSL
- Cookies
- Sesiuni

- Cel mai răspândit server web
- NCSA server, 1994
- A Patchy Web Server
- Apache Foundation
 - Apache License
 - apr, Tomcat, ant, hadoop
- Versiunea actuală:

- apache2-utils, apache2.2-bin, apache2.2-common
- apache2-doc
- apache2-mpm-... (Multi-Processing Module)
 - ...prefork – non-threaded, pre-fork (pre-creare procese daemon), stabil, compatibil
 - ...event – event-based, experimental
 - ...worker – thread-based, rapid, smaller memory footprint

- /etc/apache2/ pe Debian
- apache2.conf – main configuration file
 - include fişierele/directoarele de mai jos
- httpd.conf – compatibilitate Apache 1.x
- ports.conf
- conf.d/ - „generic snippets of statements”
- mods-available/, mods-enabled/ - configurare module
- sites-available/, sites-enabled/ - configurare site-uri (virtual hosts)
- envvars – variabile de mediu specifice
 - APACHE_RUN_USER, APACHE_RUN_GROUP, APACHE_PID_FILE
- Implicit, serverul rulează cu drepturile www-data:www-data

- Fișiere de configurare
- Executabilul asociat este apache2
- Pentru interacțiune cu serverul
 - `apache2ctl start|stop|restart|status|configtest`
 - `/etc/init.d/apache2 start|stop|restart|status`
- Jurnalizare
 - `/var/log/apache2/acces.log`
 - `/var/log/apache2/error.log`
- `a2enmod`, `a2dissmod` – activare/dezactivare module
- `a2ensite`, `a2dissite` – activare/dezactivare site-uri (virtual hosts)

- NumeDirectivă Valoare
- <Directivă> ... </Directivă>
- <IfModule ...> ... </IfModule ...>

- /etc/apache2/apache2.conf
 - <IfModule mpm_...>
 - AccessFileName .htaccess
 - ErrorLog
 - CustomLog în fișierele de site-uri (virtual hosts)
 - Include
- /etc/apache2/ports.conf
 - NameVirtualHost
 - Listen
- /etc/apache2/conf.d/security
 - ServerTokens

- apt-get install libapache2-mod-...
- /etc/apache2/mods-available/ - disponibile
- /etc/apache2/mods-enabled/ - active
- *.conf – configurare
- *.load – module efective
- a2enmod, a2dismod

- mod_dir
 - DirectoryIndex
- mod_userdir
 - UserDir public_html
 - UserDir disabled root
- mod_autoindex
 - IndexOptions

- Site = virtual host
- ServerName
 - anaconda.cs.pub.ro
- ServerAdmin
 - webmaster@anaconda.cs.pub.ro
- DocumentRoot
 - /var/www (implicit)
- CustomLog, ErrorLog
- <Directory ...> ... </Directory>

- `http://httpd.apache.org/docs/2.2/mod/mod_alias.html`
- `Alias /doc /usr/share/doc`
- `Alias /redmine /usr/local/lib/redmine`
- `Alias /git /var/cache/git`
- `AliasMatch ^/rl(|(/.*))$ /home/rl/www$1`
- `Redirect /test http://newsite/test`
- `RedirectMatch`
- `ScriptAlias /cgi-bin/ /usr/lib/cgi-bin/`
- `ScriptAliasMatch`

- <http://httpd.apache.org/docs/2.2/sections.html>
- `<Directory> ... </Directory>`
- `<Location> ... </Location>`
- `<Files> ... </Files>`
- Options (nu în `<Files>`)
 - ExecCGI, Indexes, FollowSymLinks etc.
- AllowOverride (doar pentru Directory) (.htaccess)
- Order, Allow from, Deny from (Access control) - `mod_authz_host`

- AccessFileName .htaccess
- Se recomandă să fie folosite în absența drepturilor de root
- Opțiuni controlate de AllowOverride
 - AuthConfig, Indexes, Limit, Options

- <http://httpd.apache.org/docs/2.2/howto/auth.html>
- AllowOverride AuthConfig
 - pentru fişiere .htaccess

```
<Directory ...>
```

```
AuthType basic
```

```
AuthName „restricted”
```

```
AuthUserFile /usr/local/my.htpasswd
```

```
Require valid-user
```

```
</Directory>
```

- `htpasswd [-c] /usr/local/my.htpasswd dorel`

- Despite the tons of examples and docs, mod_rewrite is voodoo. Damned cool voodoo, but still voodoo. (Brian Moore)
- RewriteEngine On
- RewriteBase /site/
- RewriteRule ^foo\.html\$ bar.html

- RewriteCond %{HTTP_HOST} ^rosedu.org\$
- RewriteRule (.*?)http://www.rosedu.org\$1 [R=permanent,L]

- ScriptAlias /cgi-bin/ /usr/lib/cgi-bin/
- Options ExecCGI
- AddHandler cgi-script .cgi
- SetHandler cgi-script
-
- mod_cgid
 - ScriptSock /var/run/cgid.sock
 - Unix MPMs only
- mod_fcgid

- `/etc/apache2/sites-available/`
- `/etc/apache2/sites-enabled/`
- Name based virtual hosting
- IP based virtual hosting
- `NameVirtualHost *:80`
- `<VirtualHost *:80> ... </VirtualHost>`

- Fișierele se creează în `/etc/apache2/sites-available/`
 - se recomandă să aibă numele de host
- `ServerName`
- `ServerAdmin`
- `DocumentRoot`
- `a2ensite`, `a2dissite`
 - se creează legături simbolice în `/etc/apache2/sites-enabled`

- <IfModule mod_ssl.c>
- SSLEngine On
- SSLCertificateFile
- SSLCertificateKeyFile
- Client Certificate Authentication
 - SSLCACertificatePath
 - SSLCACertificateFile

- Web-based Distributed Authoring and Versioning
- Set de extensii pentru HTTP pentru editarea și gestiunea fișierelor
- RFC 4918
- Writable Web
- mod_dav

```
<Location /foo>
```

```
    Dav On
```

```
    AuthType Basic
```

```
    AuthName DAV
```

```
    AuthUserFile user.passwd
```

```
    <LimitExcept GET OPTIONS>
```

```
        require user admin
```

```
    </LimitExcept>
```

```
</Location>
```

- WWW/Web
- W3C
- URL
- HTML
- HTTP, HTTPS
- server web
- browser
- Apache
- /etc/apache2/
- apache2ctl
- module Apache
- a2enmod, a2dismod
- ServerName, DocumentRoot
- Alias, Redirect, ScriptAlias
- Directory, Location, Files
- .htaccess
- mod_authz_host
- mod_rewrite
- CGI, FCGI
- virtual hosting
- A2ensite, a2dissite
- mod_ssl
- DAV

?

