

# INTERNET Z POHLEDU APLIKAČNÍHO PROGRAMÁTORA

Vlastimil Čevela

Programátorské služby Cobol, Benešova 279, 664 42 Modřice, tel. 05-4721-6183  
e-mail:cevelavl@mbox.vol.cz <http://www.vol.cz/pcweb/cevela>

## Abstrakt

Cílem tohoto příspěvku je pokus o stručný pohled na současný(03/00) stav rozvoje fenoménu Internet a možností jeho využití při tvorbě aplikací. Pokud budeme Internet chápat jako obecné multimediální prostředí s různou mírou složitosti jeho ovládání, tak bude také tvorba aplikací různě "programátorsky" náročná.

Může to být třeba návrh stránky školního časopisu s různými druhy textu, obrázky a odkazy, který např. s pomocí internetové stavebnice PCWEB zvládne iniciativní žák základní školy. Ale taky interaktivní grafika, komunikující s databází, která vyžaduje znalost objektových modelů, stylů, komponent, skriptovacích či programovacích jazyků, anglicky komunikujících vývojových nástrojů, atd atd. V tom druhém případě jsou však asi nejdůležitější pevné nervy a jistý nadhled, aby se člověk neutopil v záplavě informací a nových technologií, které se na něj průběžně valí.

## 1. Základní fakta kolem WWW a HTML(XML)

Původní myšlenky, WWW (World Wide Web) projektu, které se prakticky dodnes uplatňují byly velice jednoduché:

- a) Jako HTML (Hypertext Markup Language) stránky nejsou na serveru uloženy a po síti posílány kompletně dotvořené výstupy, ale pouze prostý text, doplněný smluvenými značkami, se kterými umí pracovat univerzální program (prohlížeč) jako klient.
- b) V rámci čistě textových stránek mohou být uvedeny URL (Uniform Resource Locator) adresy na samostatné binární soubory (obrázky, zvuky, video), i na spustitelné programy (skripty, aplety a komponenty). Ze všech těchto elementů jako ze stavebnice pak prohlížeč seskládá výslednou prezentaci.
- c) Zobrazené stránky mohou dále obsahovat zvýrazněné hypertextové, t.j. "na ukázání v textu reagující" odkazy na další URL adresy. Po výběru tlačítkem myši pak prohlížeč požádá příslušný server o načtení jiné stránky.

Protože URL adresy uložení všech stránek v celosvětové síti se řídí jednotnými pravidly (domény pro jednotlivé státy s centrální registrací připojených serverů a převodníkem na 4-skupinové číselné IP adresy), může kterýkoliv uživatel jako klient pomocí prohlížeče na svém počítači prakticky neomezeně přecházet mezi všemi stránkami WWW.

Nejnovější doporučení W3C (World Wide Web Consortium) pak oficiálně specifikují Dynamické HTML s objektovým modelem dokumentu DOM (Document Object Model) a přechod na XML (eXtensible Markup Language). Zdá se, že v podobě XML roste nový, tentokrát asi už přímo průmyslový standard, který nejenom sjednocuje a zpřisňuje dosavadní pravidla HTML, ale umožňuje na WWW stránky umístit též informaci o významu jednotlivých částí struktury textů - to by mělo být velice důležité pro zvýšení kvality výsledků při vyhledávání.

## 2. Internet - Intranet - Extranet

Obecně lze říci, že ve všech případech z nadpisu této kapitoly se jedná o prakticky shodné základní, řekněme "internetové" technologie. Rozdíl je pouze v tom, že INTERNET je veřejně přístupná světová síť, INTRANET označuje interně řízenou a ohraničenou síť uvnitř organizace či školy a EXTRANET pak distribuované části Intranetu, propojené přes Internet.

Prakticky všude se shodně používá TCP/IP (Transmission Control Protocol / Internet Protocol) pro přepravu dat mezi partnery, HTTP (Hypertext Transfer Protocol) pro dialog mezi WWW klientem a serverem, URL + IP a DNS (Domain Name System) pro specifikování adres dokumentů/objektů, CGI (Common Gateway Interface) pro předávání informací z vyplněných formulářů na server, MIME (Multipart Internet Mail Extensions) pro přenos multimediálních dat pomocí poštovních standardů Internetu, či pracuje s objekty, skripty atd.

Rozdíl je především v organizaci věcného obsahu a způsobu zabezpečení oprávněnosti přístupů. Z hlediska organizačního se v uvedených souvislostech velice naléhavě objevuje potřeba průběžné kvalifikované správy věcného obsahu všech systémů na bázi internetových technologií. Bohužel tato potřeba není ani na úrovni našich státních orgánů, ani ve většině organizací zatím dostatečně intenzivně vnímána.

Pro aplikační programování nabízí Internet obrovské možnosti úspor nákladů, pracnosti a zvýšení produktivity. Je tady totiž k dispozici proti dřívějšímu nesouměřitelná, celosvětově jednotná, levná, z velké části prakticky hotová a obecně použitelná technologie pro přenos, zpracování a prezentaci libovolných dat. Zdá se však, že ani v odborných kruzích především ze strany střední a starší analyticko-programátorské generace nejsou nabízené možnosti oceňovány a využívány tak, jak by si zasluhovaly.

Nevěnovat pozornost Internetu je pro stát, natožpak pro specialisty z oboru, určitě velká škoda, protože nikdy dříve v historii rozvoje informačních technologií se podobně významný a přitom levný technologický skok zatím nevyskytl. V dnešní době se silnou konkurencí a globalizací navíc už i u nás platí Nerudovo "kdo chvíli stál, už stojí opodál".

## 3. Programování na straně klienta

Prohlížeč WWW - klient, si dle požadavku uživatele z příslušného serveru zkopíruje požadovanou stránku v jazyku HTML s příp. dalšími vnořenými objekty (obrázky atd), dočasně si ji uloží do pracovní oblasti na svém lokálním disku a prezentuje seskládaný výsledek. Základem programování na straně klienta je tedy vytvoření HTML stránky, tj. běžného textového souboru se značkami (tagy), dle kterých pak prohlížeč realizuje funkce, potřebné k prezentaci.

Za praktický standard se dají považovat 2 světově nejrozšířenější řady prohlížečů - Netscape Navigator a MS Internet Explorer. Pokud nemají některé specializovanější funkce (např. přehrávání videa, práce s komponentami a pod.) přímo v sobě zabudovány, bývají pro jejich realizaci zpravidla k dispozici tzv. "plug-in" (zásuvné, výměnné, doplňující) moduly.

Elementy ve struktuře HTML stránky/dokumentu můžeme rozdělit na několik částí. Především tedy text s vlastní informační hodnotou, pak pokyny k formátování při jeho zobrazení a reference na další objekty, potřebné k prezentaci. Součástí stránky může také být prohlížečem interpretovaný zdrojový text programu např. v jazyku MS-VBScript či JavaScript, anebo odkaz na již zkompilevanou programovou komponentu. Skripty a programové komponenty pak mohou řídit prezentaci stránek prohlížečem na základě

událostí, tj. pomocí interakce s uživatelem, jako je např. práce s myší, výpočty/grafika dle zadávaných hodnot a pod.

Předem připravenou komponentou může být prakticky jakýkoliv programový objekt, který dodržuje pravidla technologií COM (Component Object Model), Java nebo CORBA (Common Object Request Broker Architecture), tj. třeba v jazyku C++, OO Cobol či jiném. Bezpečnostní zásadou však je, že komponenty, programované pro práci na počítači klienta obecně nesmějí mít možnost zápisu na jeho privátní diskové a paměťové prostory.

Pro ovládání objektů slouží kromě výše uvedených skriptovacích jazyků též Aplety, které fungují tak, že se při každém načtení stránky klientem načítají znovu a komponenty COM (MS ActiveX). Ty jsou z centrální knihovny na WWW stahovány na klientský počítač jen jednou, zaregistrují se u něj, a následně jsou opakovaně využívány. V uvedených souvislostech hrají důležitou roli tzv. autorizační certifikáty technologie Authenticode, která spočívá v nezávislém ověření původu určitého elementu, se kterým hodlá klient dále pracovat. Je dobré vědět, že certifikát potvrzuje původ a nikoliv správnost - za tu zodpovídá výrobce příslušného objektu, a taky že dle současných znalostí takový elektronický podpis není možné padělat.

Existují ještě další dvě možnosti, kdy stránka, stahovaná ze serveru může zapsat data na počítač klienta. Jednou z nich jsou skrytá pole HTML, kdy server při posílání stránky s formulářem klientovi do ní zapíše nějaké informace, které při návratu téhož formuláře od klienta pak může vyhodnotit. Pokud si ale uživatel uvedený formulář neuloží a před další komunikací s příslušným serverem počítač vypne, nebude tato zpětná vazba fungovat. Druhou možností jsou tzv. "cookies" (koláčky), kdy stránka načítaná ze serveru si ve vyhrazeném adresáři na počítači klienta může uložit malý textový soubor s identifikací serveru a uživatele, obsahující příslušné informace a dobu platnosti a při opakovaném kontaktu jej může číst. Uživatel však ve svém nastavení prohlížeče nemusí zápis "cookies" povolit.

Jazyk HTML ve svých dosavadních verzích, ve kterých je v současné době prezentována většina stránek na WWW, je zaměřen na problém "jak zobrazovat informace". Znamená to, že má nástroje (styly) pro různé druhy písma, zvýraznění a barev, na tvorbu výřtů definic či seznamů, sestavování tabulek, příp. stránky z více oken (rámy), odstavce, odkazy, formuláře, vkládání obrázků atd. V poslední době se pak začíná prosazovat jeho dynamické rozšíření a objektové modely HTML dokumentu i prohlížeče a novému standardu XML, oficiálně ze strany W3C doporučenému v lednu 2000 se předpovídá velká budoucnost.

#### **4. Programování na straně serveru**

Viditelná komunikační rozhraní mezi klientem tj. prohlížečem na počítači uživatele a serverem jsou dvě - hypertextový odkaz nebo konkrétně napsané zadání URL požadované stránky a odeslání vyplněného formuláře. V obou případech je výsledkem činnosti serveru vrácení nové stránky klientovi. V prvním zná přímo její URL adresu, zatímco ve druhém pouze jméno programu/skriptu, který má spustit. Teprve tento program/skript pak na základě vyplněných vstupních dat ve formuláři a příp. skrytých polí zařídí získání, vytvoření či doplnění nové stránky.

Obecně může server klientovi poslat požadovanou stránku tak jak je, anebo na přímo na ní mohou být pokyny k provedení úprav či výpočtů. Mohou to být buď opět zdrojové texty skriptů, nebo předkompilované komponenty - např. ASP (Active Server Pages) a další. Server pak ještě před odesláním stránky příslušné pokyny realizuje.

Jako odezvu na formulář pak může server (podle typu a dodavatele) spouštět opět buď skripty, anebo programy, které přebírají data z formuláře prostřednictvím rozhraní CGI. Takto lze spouštět skripty operačního systému (Unix), speciálních skriptovacích jazyků (Perl, PHP), anebo zkompilevané programy, napsané v některém z jazyků, které CGI umí (C++, Cobol). Tím, že je na serveru takto k dispozici vlastně možnost běžného programování, tak lze samozřejmě pracovat se soubory, komunikovat s databázemi a realizovat libovolně složité výpočty.

Jsou zde však ještě dva důležité problémy - bezpečnost a průchodnost. Z hlediska bezpečnosti by mělo být zajištěno, aby nebylo možné se ze strany klienta neautorizovaně dostat k ovládání dat a programů na serveru. S tím pak souvisí potřebná omezení, která musí být na serverové programy i na stránky s aktivními prvky a skripty kladena.

Při větším množství požadavků na server je třeba natahovat do paměti a spouštět nebo interpretovat mnoho programů, což je velmi náročné. Proto se od původního spouštění samostatných CGI skriptů přešlo na využívání dynamických knihoven, které jsou na serveru trvale v paměti (např. DLL pro ISAPI), anebo přímo přikompilovány do serveru (např. modul PHP pro server Apache). To už ale spolu s ošetřením vláken a obecně řízením práce na serveru není předmětem tohoto příspěvku.

## 5. Další možnosti využití internetové technologie

Oblasti, ve kterých se úspěšně a ve velkém stylu uplatňuje WWW služba Internetu jsou předmětem mnoha článků, úvah i TV relací. V tomto příspěvku se zaměříme na několik střídavějších, spíše doplňujících - ale přesto o nic méně užitečných aplikačních možností.

Tak především je dobré vědět, že jeden z celosvětově nejužívanějších internetových serverů Linux - Apache, který poskytuje plnohodnotnou klient/server interakci, je k dispozici zadarmo. Pro jednostranné prohlížení WWW stránek, umístěných na serveru lokální sítě pak je protokol TCP/IP běžnou výbavou a na každou pracovní stanici s grafikou pod Windows nebo Linux lze rovněž zdarma umístit prohlížeč MS Internet Explorer nebo Netscape Navigator.

Interaktivně jednosměrný Intranet, tj. prezentace libovolných stránek ve školní nebo podnikové síti (bez formulářů a problémů s programováním na serveru) je tedy možno kdekoliv realizovat prakticky okamžitě a bez zvýšených nároků na programátorskou správu sítě. Stačí pro ně zřídit sdílený prezentační prostor a pověřit zodpovědnou osobu správou věcného obsahu. Užitečný hypertextový informační systém, který ovládá různé formáty dat je možno provozovat i na vlastním osobním PC. Stačí k tomu mít nainstalovaný prohlížeč a pár řídicích stránek s odkazy, vytvořených pomocí jakéhokoliv nástroje pro psaní HTML, anebo prostě textovým editorem.

WWW prohlížečem je možno pomocí hypertextových odkazů prezentovat nejenom stránky ".htm"/".html", ale např. též soubory elektronické pošty ".eml" včetně příloh a především obyčejné ASCII textové soubory ".txt", které se zobrazí jako pevně předformátovaný text. Když to trochu domyslíme třeba pro podmínky podnikového výpočetního střediska tak zjistíme, že tímto způsobem můžeme doslova zadarmo, tj. bez jakýchkoliv dodatečných HW a SW nákladů a bez jakéhokoliv programování prakticky okamžitě standardně zpřístupnit k prohlížení a k příp. individuálním tiskům všechny výstupní sestavy ze sálového dávkového zpracování dat.

Prohlížeč hravě zvládne i prezentaci 160 sloupců, pouze sestavy s velkým počtem stránek bude potřeba rozdělit na několik menších. A vzato do důsledků, nemusíme ty sestavy ani umisťovat na síť, ale můžeme je na jednotlivá izolovaná pracoviště předávat třeba na

disketě. Pro jejich HW/SW vybavení přitom stačí i PC 386/486 s Windows 3.x a prohlížeč MSIE 3.02a. Někomu se tyto úvahy budou zdát primitivní, ale existuje hodně organizací i škol, které nemají na velkorysé inovace a tak jim nezbývá, než využívat existující technologické zdroje na maximum. A celkem vzato to z ekonomického hlediska ani možná není tak úplně hloupé ?!

### **Související literatura a odborné podklady**

1. Baranovič R., Moravčíková L., Šnajder L., Internet pro střední školy, Computer Press Praha 1999 (všeobecný a komplexní přehled !)
2. Kosek Jiří, HTML - Tvorba dokonalých www stránek - Podrobný průvodce, Grada Publishing Praha 1998 (metodicky velice dobrý návod !)
3. Pirouz Raymond, HTML kouzla na webu - Praktické postupy pro vytváření atraktivních webových stránek, Unis Publishing Brno 1998 (řada zajímavých a vtipných nápadů !)
4. Hlavenka J., Sedlár R., Holčík T., Šebesta M., Botík R., Vytváříme www stránky a spravujeme moderní WEB SITE, Computer Press Praha 1997 (celkový přehled, základy ASP !)
5. Mikle Pavel, Dynamické HTML - Referenční příručka, Unis Publishing Brno 1997 (nejvodnější referenční podklad !!!)
6. Hatfield Bill, Visual Basic Script - Tvorba dokonalých stránek, Grada Publishing Praha 1997
7. Cornell Gary, Microsoft Visual Basic Scripting - Příručka programátora, Computer Press Brno 1999
8. Purcell L., Mara M. J., JavaScript - Tvorba dokonalých www stránek - podrobný průvodce začínajícího uživatele, Grada Publishing Praha 1998
9. Kosek Jiří, PHP - Tvorba interaktivních internetových aplikací - podrobný průvodce, Grada Publishing Praha 1999
10. Haggard Mary, Příručka tonoucího webmastera - Jak plánovat, vytvářet a spravovat kvalitní a úspěšný web, Computer Press Brno 1999
11. Hlavenka Jiří, Dělejte byznis na Internetu - Jak využívat Internet k prospěchu firmy i jednotlivce, Computer Press Praha 1999
12. Greer Tyson, Intranety - Principy a praxe - Průvodce nasazením sítí Internetového typu v podniku, Computer Press Brno 1999
13. "www.kosek.cz" - Téměř vše o WWW, PHP, XML (!!!)
14. "www.vol.cz/pcweb" - Internetová stavebnice PCWEB (nástroj pro tvorbu HTML stránek !)
15. "w3c.org" - Činnost a oficiální doporučení World Wide Web Consortium (!!!)
16. "msdn.microsoft.com/workshop/server/feature/security.asp" - Authentication and Security for Internet Developers (!)
17. "www.dhtmlzone.com" - Dynamic HTML (zajímavé ošetření událostí!)