

## TECHNOLOGIE MULTIMEDIÁLNÍHO PUBLIKOVÁNÍ NA DVD ZKUŠENOSTI Z PROJEKTU 10@FI

Petr Sojka

### Abstrakt

*Publikování a šíření multimediálních a výukových materiálů na nosičích DVD je stále častěji. U každého většího publikačního projektu však prefabrikovaná řešení obvykle nestačí - je třeba řešení ušít na míru. Článek diskutuje technologie, formáty a postupy vybrané a odzkoušené při tvorbě DVD připraveného k desátému výročí Fakulty informatiky MU s názvem 10@FI. Čtenář se seznámí s průběhem přípravy DVD od prvotní myšlenky a návrhu až po konečné vytvoření (GNU/Linux) bootovatelného obrazu DVD a jeho vylisování. Celý projekt vznikl převážně pod OS Linux s pomocí programů s otevřeným zdrojovým kódem, při širokém využití XML technologie a standardů W3C (desítky spolupracujících, ke stovce spoluautorů, tisíce provázaných souborů, stovky obrázků a fotek, desítky minut původních filmů).*

### Klíčová slova

*DVD, multimédia, elektronické publikování, XML, XSLT, DocBook*

### 1 Úvod

Multimediální prezentace na médiích CD a DVD jsou dnes standardním prostředkem šíření informací. Kvalitní příprava velkého multimediálního díla na DVD je obrovský projekt, jehož příprava a realizace zahrnuje široké spektrum činnosti několika desítek specializovaných profesí.

Po poměrně dobrých zkušenostech s vysoce svépomocnou (studenti a pedagogové FI) přípravou CD *Všech 5 pohromadě* [5] v roce 1999 Fakulta informatiky zvažovala možnosti, jak se zviditelnit u příležitosti svého letošního (2004) desátého výročí. Vývoj informačních technologií pokročil mílovými kroky kupředu a CD nosiče jsou nahrazovány DVD nosiči. Vystala otázka, zda si prezentační DVD nechat vyrobit nějakou renomovanou firmou, nebo si akademičtí informatici zase zkusí prakticky zpracovat informace informačními technologiemi sami. Pevní varianta by dozajista nevypovídala o FI zdaleka tolik, jako druhá. Jelikož byl učiněn pokus vypsát zaměření předmětu PV113 *Softwarové elektronické publikace* tematicky zaměřené na praktickou přepravu DVD, nestálo realizaci projektu sedmnácti zapsanými studenty pod vederním tří pedagogů

nic v cestě: jeden z těchto studentů pod vederním jednoho z pedagogů o tomto projektu tímto článkem referuje.

## 2 Hile a infrastruktura projektu

### 2.1 Hile

Úvodem byly vytčeny tyto hile: DVD má zmapovat desetiletou existenci Fakulty informatiky Masarykovy univerzity, sloužit jako archiv studijních materiálů, učebních textů, má prezentovat jednotlivé výzkumné laboratoře i bila studentů, zejména pak dát plastický. obrázek o škole zájemcům o studium a začínajícím studentům. Rozhodlo se o nákladu 3000 ks datových DVD typu DVD5 (4,7GB), začali se shánět sponzoři, ale hlavně se začala řešit technologie přípravy DVD a softwarová podpora redakční práce týmu.

### 2.2 Komunikační prostředky

Je všeobecně známo, že u větších softwarových projektů je klíčové definovat *rozhraní*, a to jak *datová*, tak *komunikační*, a ta pak důsledně používat, resp. jejich užívání vynucovat. Jako hlavní komunikační nástroj kromě elektronické pošty (tým je dostupný na aliasu 10@fi.muni.cz) pro kolaborativní sdílení průběžných informací a diskuse přes web byl zvolen systém WIKI (<http://wiki.org>). WIKI umožňuje snadno měnit obsah dokumentů, udržovat informace o změnách a uživatelích. Tým se třináct týdnů scházel jednou týdně celý a kromě toho se dle potřeby scházely jednotlivé podtýmy.

### 2.3 Infrastruktura

Jako zázemí pro tvorbu sloužily tyto tři laboratoře: SITOLA (Laboratoř pokročilých síťových technologií), LEMMA (Laboratoř elektronických a multimediálních aplikací) a hlavně NLP (Laboratoř zpracování přirozeného jazyka), ve které je umístěn hlavní pracovní server. Na projektu pracovala dvacítká stálých členů plus nespočet externistů.

## 3 Textová část

### 3.1 Konverze z CD Všech 5 pohromadě

Vzor pro naše DVD bylo CD Všech 5 pohromadě a protože jsme na DVD chtěli zachovat původní data a dále je rozšiřovat, museli jsme dokumenty, které byly ve speciálně navrženém metajazyce, konvertovat na námi zvolené DocBookovské XML. Díky tomu, že tvůrci CD již v roce 1999 navrhovali dokumenty tak, aby byl obsah oddělený od formy, nebylo velkým problémem konverzi uskutečnit.

### 3.2 DocBook

Jako nejvhodnější koncový formát dokumentů pro DVD jsme zvolili XHTML. Bylo by však neefektivní vytvářet přímo XHTML dokumenty (např. pro tisk za účelem korektury je vhodnější PDF), proto jsme použili modifikovaný DocBook [6]. Technologie multimediálního publikování na DVD: zkušenosti z projektu 10@FI

Díky DocBooku stačí vytvářet jeden XML soubor, který pak můžeme dále konvertovat do různých formátů (v našem případě XHTML, PDF pro korektory, nebo PostScript). DocBook je formát XML dokumentu a popisuje tedy dokumenty na sémantické úrovni, ne na vizuální, to znamená, že se píše obsah, ale už se nepíše, jak se má daný obsah zobrazit.

### 3.3 Transformace dokumentů

Abychom z XML souboru dostali XHTML, poslouží XSL transformace [2]. XSLT procesor si vezme XML, aplikuje na něj pravidla ze stylu a výsledek uloží v XHTML. Pro vygenerování PDF dokumentu se XML transformuje (XSLT procesorem) na FO [1] soubor, z kterého se FO procesorem vygeneruje požadovaný PDF dokument. Sám XSLT procesor totiž PDF (binární soubor) vygenerovat neumí. Jako FO procesor jsme používali FOP.

K převodu do XHTML posloužil XSLT procesor XSLTPROC, k validaci XML souboru XMLLINT.

DocBook také plně podporuje schopnosti různých výstupních formátů (odkazy, rejstříky ... ), což je velmi výhodné.

Při zpětném pohledu se zdá, že modifikace DocBookového DTD nebyla nejlepší volba. Mnohem efektivnější je vytvořit vlastní DTD pro dokumenty, vlastní XSLT šablony pro převod do XHTML a DocBooku. šablona pro převod do DocBooku je důležitá kvůli dalšímu převodu do PDF, jelikož vytvořit si XSLT šablonu pro převod do PDF (přesněji řečeno do FO souboru), není zrovna triviální záležitost.

### 3.4 Metainformace

DVD si zachovalo původní kostru z CD *Všech 5 pohromadě*, tím pádem bylo vhodné vytvořit dva druhy metainformací pro XML soubory - jiné pro kořenové XML kapitol a jiné pro samotné dokumenty.

Vzor DTD pro kořenové XML soubory:

```
<!-- metainformace, indexy jednotlivých kapitol -->
<!ELEMENT kapitola (nazev, (kapitola*, dokument*)*)>
  <!ATTLIST kapitola
    video CDATA #IMPLIED
    soubor CDATA #IMPLIED
    prekladat (ano | ne) #IMPLIED
    autorska-prava (ano | ne) #IMPLIED
    korektura (ano | ne) #IMPLIED
    kotva CDATA #IMPLIED
    popis CDATA #IMPLIED>
<!ELEMENT nazev (#PCDATA)>
<!ELEMENT dokument (spatny-odkaz*)>
  <!ATTLIST dokument
```

```
soubor CDATA #REQUIRED
prekladat (ano | ne) #IMPLIED
autorska-prava (ano | ne) #IMPLIED
korektura (ano | ne) #IMPLIED
kotva CDATA #IMPLIED
popis CDATA #IMPLIED
dontshow (ano|ne) #IMPLIED>
```

```
<!ELEMENT spatny-odkaz (#PCDATA)>
```

Vzor DTD pro dokumenty:

```
<!-- metainformace umistene v~hlavice dokumentu -->
```

```
<!ELEMENT autor EMPTY>
  ATTLIST autor id CDATA #REQUIRED>
<!ELEMENT zdroj (#PCDATA)>
<!ELEMENT odpovedna-osoba EMPTY>
  <!ATTLIST odpovedna-osoba id CDATA #REQUIRED>
<!ELEMENT datum EMPTY>
  <!ATTLIST datum od CDATA #REQUIRED
                do CDATA #IMPLIED>
<!ELEMENT poznamka (#PCDATA)>
```

Pro jednoduchý zápis jmen je vhodné udržovat seznam osob a odkazovat na něj pomocí id osoby.

Ukázka DTD metalide.dtd:

```
<!ELEMENT lide (osoba*)>
  <!ELEMENT osoba (jmeno, prijmeni, titul?, foto?, popisek?)>
  <!ATTLIST osoba id ID #REQUIRED>
    <!ELEMENT jmeno (#PCDATA)>
    <!ELEMENT prijmeni (#PCDATA)>
    <!ELEMENT titul EMPTY>
    <!ATTLIST titul pred CDATA #IMPLIED
                  za CDATA #IMPLIED>
    <!ELEMENT foto (#PCDATA)>
    <!ELEMENT popisek (#PCDATA)>
```

a seznamu osob lide.xml:

```
<?xml version="1.0" encoding="iso-8859-2"?>
<!DOCTYPE lide SYSTEM "metalide.dtd">

<lide>
  <osoba id="LMatyska">
    <jmeno>Luděk</jmeno>
    <prijmeni>Matyska</prijmeni>
    <titul pred="Doc. RNDr." za="CSc." />
    <foto>f10://602342354567687978</foto>
```

```
<popisek>D`Ikan FI</popisek>
</osoba>
</lide/>
```

Podobně metainformace jsme vytvořili i pro videa a fotky. Do metainformací se ukládají informace o stavu dokumentu (korektury, autorská práva, jedinečné id objektu. . .). Díky metainformacím se snadno vytvářejí rejstříky (osob, videa, fotek a tématický rejstřík).

### **3.5 Kontrola odkazů**

Kontrola odkazů je realizována v jazyce Java, která má velmi dobrou podporu jazyka XML. Bohužel velkou nevýhodou Javy je její pomalost. Kontrola odkazů se provádí rekurzivním průchodem stromové struktury XML souborů od kořene. Lze zaznamenat počet odkazovaných souborů, počet přebytečných souborů, počet chybějících souborů a dokonce rozeznávat chybějící/nadbytečné soubory podle typu (obrázky, video, dokumenty. . .).

## **4 Grafická část**

### **4.1 Design**

Vzhled XHTML stránek zajišťují kaskádové styly (CSS) a díky javascriptu lze vzhled jednoduše přepínat (pouze se mění soubor s kaskádovými styly) - čtenář má na výběr z několika vzhledů. Jelikož je forma dokumentu skriktně oddělena od obsahu (součástí XHTML kódu nejsou žádné tagy určující vzhled), má úprava jednoho CSS souboru vliv na vzhled všech XHTML stránek na DVD.

### **4.2 Logotyp, obal a potisk**

Již od začátku jsme promýšleli ucelený a jednotný vzhled celého DVD, kam samozřejmě spadá i logotyp, obal a potisk. Bylo navrženo více grafických variací od různých autorů a více názvů pro DVD. O vítězném návrhu se demokraticky hlasovalo. Výsledný vzhled logotypu a návrh názvů pro DVD jsou na obrázku 1.

## **5 Multimediální část**

### **5.1 Videomedailonky**

Díky velikosti DVD mohl multimediální podtým vytvořit videomedailonky jednotlivých laboratoří fakulty v délce několika desítek minut. Medailonků je devět. Jsou to kompletně autorská díla studentů od námětu, scénáře, režie, ozvučení a střihu až po finální produkci. Jeden ze studentů navíc vymodeloval a "oživil" virtuální hlasatelku, která některé z medailonků uvádí. Dalším zpestřením jsou ukázky z každoročně pořádaného filmového festivalu studentských filmů. Na DVD se tudíž nachází celkem přes hodinu autorského videa.

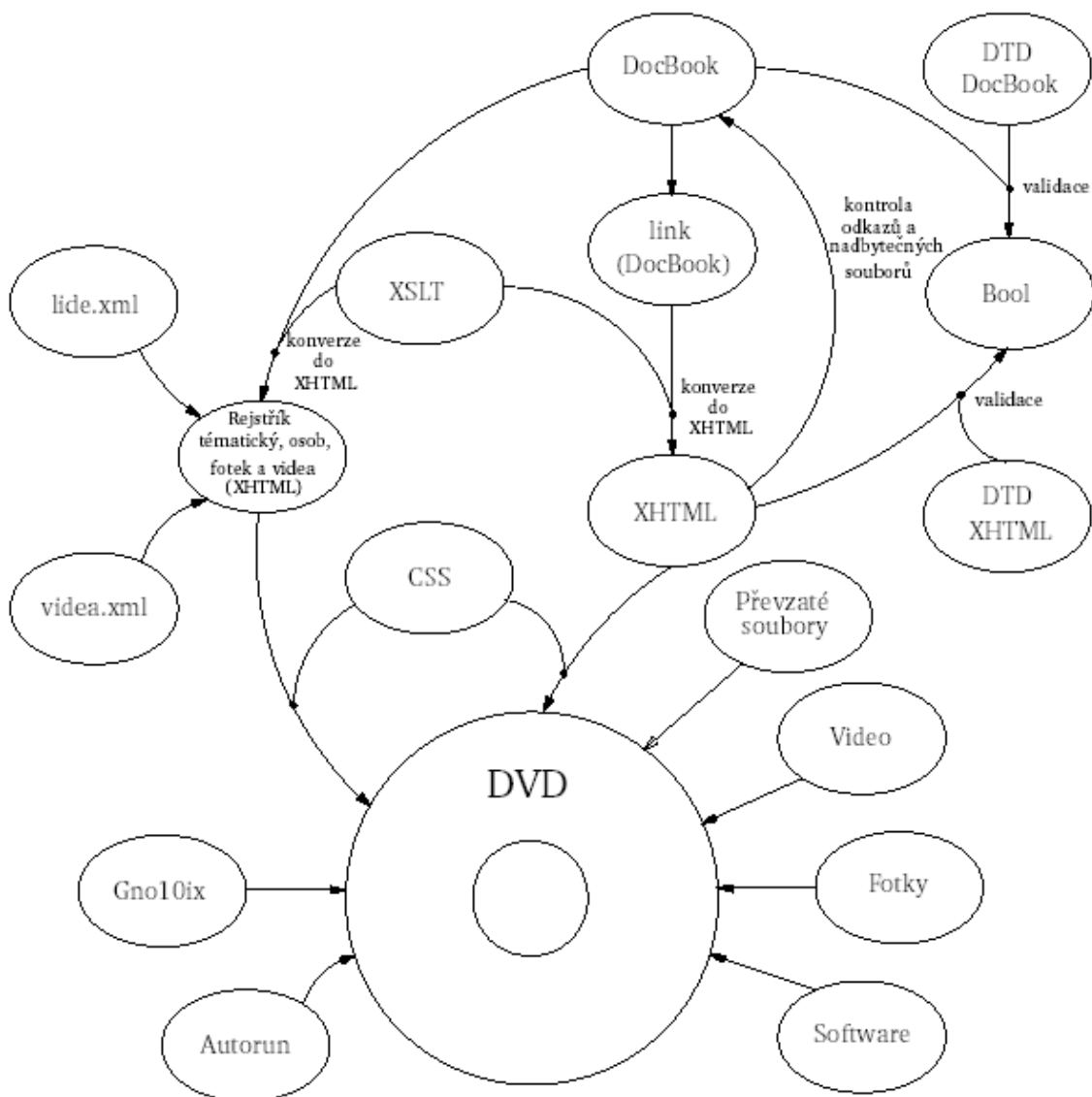
Jako komprimační kodeky byly použity XVID verze 1.0 pro video a MP3 pro zvukovou stopu. XviD kodek je pod licencí GPL a licenční podmínky u nejznámějšího audio kodeku MP3 jsou v projektu splněny. U původně zvažovaného OGG VORBIS se objevily problémy s přehráváním v komerčních rozšířených přehrávačích videa. Stříh videa probíhal na profesionálním stříhacím systému AVID XPRESS PRO. Pro finální zkompletování zvuku s videem, deinterlacing a kompresi dokonale postačoval volně šířený program VIRTUAL DUB.



Obrázek 1. Logotyp projektu DVD 10@FI a návrhy názvu DVD

## 6 Generování DVD

Výsledný proces vygenerování ISO obrazu DVD se skládá z téměř desítky kroků, které jsou znázorněny na obrázku 2. Zdrojový XML soubor se zvaliduje proti DTD DocBooku. Pokud je validní, nahradí se v něm symbolické odkazy za skutečné a soubor se uloží s příponou link (stále se jedná o validní DocBookové XML). Zároveň se vytvoří rejstříky osob, fotek, videí a rejstřík tematický. Pak se link-souborů a rejstříků vygenerují XSLT procesorem XHTML soubory, které se opět podrobí validaci (samozřejmě, že již ne oproti DTD DocBooku, ale oproti DTD XHTML). Nad XHTML soubory se provede kontrola odkazů a nadbytečných souborů - chybějící a nadbytečné soubory se vypíší na obrazovku, editor musí doopravit ručně. Výsledné XHTML soubory spolu s fotkami, videi, mirrory stažených dokumentů a stránek, autorunem, softwarem (viz oddíl 7.4) a Gno10ixem (viz oddíl 7.2) jsou již připraveny k prezentaci na DVD. Vše je automatizováno s pomocí GNU MAKE.



Obrázek 2. Tok dat při generování DVD

## 7 Ostatní

### 7.1 Správa verzí

Projekty, na nichž pracuje mnoho lidí, je potřeba nějak zabezpečit proti vzájemnému přepisování souborů, zajistit možnost navracení k předchozím verzím atd. My jsme volili z osvědčeného CVS a v té době nového SUBVERSION. CVS má, jak je známo, dva hlavní nedostatky: 1) neumí přejmenovat soubor bez toho, aniž bychom ztratili informaci o logu souboru a 2) binární soubory přidávané do repozitáře musí být ručně označeny speciálním parametrem, jinak může dojít k jejich poškození. My se proto rozhodli použít SUBVERSION, který plně nahrazuje CVS, a má tyto problémy řešit. Bohužel jsme během nasazení SUBVERSION narazili na problém s náhodným padáním

verzovacího systému (mylně přisuzovaný tomu, že běžel na víceprocesorovém stroji) a jelikož jsme byli v časové tísní a nebyl čas problém řešit, vrátili jsme se k CVS.

SUBVERSION nám padalo z důvodu špatně nastavené masky (022 místo 002), proto se musí binární soubory SVN a SVNSENVE nahradit skripty, které nejdřív masku nastaví a pak spustí samotné binární SVN či SVNSENVE.

System správy verzí neřešil pouze výše zmiňované, ale zajišťoval přes commit skripty antivirovou kontrolu posílaných souborů a kontrolu validity XML souborů.

Více o verzovacích systémech čtenář nalezne v samostatném článku v tomto sborníku [7].

## **7.2 Bootovatelnost média**

Pro použití DVD není zapotřebí mít nainstalován operační systém, DVD je totiž bootovatelné. Je založeno na linuxové live distribuci GNOPPIX (a ta dále na DEBIAN GNU/LINUX), pojmenované GNO10IX. V distribuci je předinstalované grafické pracovní prostředí GNOME a veškeré potřebné programové vybavení. Naše live distribuce zabírá cca 400MB prostoru, to není málo, což vyvažuje uživatelskou přítulností a nejnovějšími verzemi programů v ní obsažených. Více čtenář nalezne v článku [3].

## **7.3 Betatestování**

Při takto rozsáhlém projektu je důležité celý projekt řádně otestovat, nejlépe někým, kdo nemá s výrobou samotného DVD nic společného a takový člověk najde často nejvíce chyb. Proto byly průběžně páleny betaverze DVD na DVD-RW a půjčovány dobrovolníkům k testování. Tímto způsobem bylo zjištěno mnoho chyb, na které bychom přicházeli jen velice stěží nebo vůbec. Navíc byla vždy aktuální verze DVD dostupná i přes WWW, takže se na testování mohlo podílet mnohem více lidí, než jen těch několik, kteří měli půjčenou betaverzi DVD.

## **7.4 Software na DVD**

Protože ne každý má ve svém počítači nainstalovaný potřebný software, umístili jsme na DVD vše, co je k práci s DVD třeba. Každý software je na DVD uložen ve verzi pro MS WINDOWS a GNU/LINUX, ať už jde o oblíbený prohlížeč MOZILLA FIREFOX, ADOBE READER, nebo kodeky XVID pro video a OGG VORBIS pro zvuk. Tyto programy jsou volně šířitelné a s volně dostupnými zdrojovými kódy (vyjma ADOBE READER).

Uživatelům systému MS WINDOWS automatizuje detekci software autorun obsažený na DVD, který se snaží zjistit nainstalované součásti systému a pokud nějaký potřebný software chybí, nebo je v systému zastaralá verze, nabídne uživateli instalaci software potřebného pro práci s DVD.



## 8 Shrnutí a zkušenosti

Příprava finálního produktu jako součást výuky u takovýchto velkých projektův akademickém prostředí s sebou přináší problém s dodržováním termínů, na druhou stranu se ovšem výsledný produkt téměř rovná komerčním s využitím minimálního rozpočtu. Proto bylo důležité dělat časté týmové schůze, kontroly kvality a adekvátní přidělování práce studentům podle jejich dispozic. Nezbytné bylo také vytvoření prostředí a atmosféry pro komunikaci mezi jádrem realizačního týmu a akademickou obcí FI. Bylo by nemyslitelné DVD vytvořit bez vysokého stupně automatizace pomocí Makefile. Dosažení potřebné kvality bylo možné díky speciálně vyvinutým automatizovaným i ručním procedurám (kontroly odkazů, konverze dokumentů, . . .).

### 8.1 Poděkování

Jak naznačeno, na projektu se podílelo autorsky kolem stovky osob, produkčně přes dvacet. Je mimo možnosti tohoto reportu vyjmenovat zásluhy každého zvlášť..

### Reference

1. Sharon Adler, Anders Berglund, Jeff Caruso, Stephen Deach, Tony Graham, Paul Grosso, Eduardo Gutentag, Alex Milowski, Scott Parnell, Jeremy Richman, a Steve Zilles. Extensible Stylesheet Language (XSL), 2001.  
<http://www.w3.org/TR/xsl>.
2. James Clark. XSL Transformations (XSLT), 1999. <http://www.w3.org/TR/xslt>.
3. Marcel Kolaja. Live distribuce aneb Linux na CD/DVD. V Sborník SLT 2004, Znojmo, 2004. s. 203-212. Konvoj.
4. Petr Vopálenský, Petr Sojka. Multimediální publikování na DVD: projekt 10@FI. V Sborník SLT 2004, Znojmo, 2004. s. 21-29. Konvoj.
5. Všech pět pohromadě. CD-ROM, <http://nlp.fi.muni.cz/projekty/vsech5/>, září 1999.
6. Norman Walsh a Leonard Mueller. DocBook: The Definitive Guide. O'Reilly & Associates, 1999.
7. Milan Zamazal. GNU arch, systém pro správu verzí. V Sborník SLT 2004, Znojmo, 2004. Konvoj.

### Petr Sojka

Masarykova univerzita v Brně, Fakulta informatiky  
Botanická 68a, 602 00 Brno  
e-mail: [sojka@fi.muni.cz](mailto:sojka@fi.muni.cz)