

Přehled open source systémů pro správu obsahu (CMS)

Bc. Petr Vašíček

20. prosince 2005

Úvod

Systémy pro správu obsahu jsou počítačové systémy sloužící k vytvoření, organizaci a řízení obsahu, ať už se jedná o čistě webovou prezentaci či přímo digitální knihovnu. Nejčastěji se jedná o webové aplikace, ale nemusí tomu tak být vždy, někdy je pro vytváření a úpravu obsahu třeba použít speciální klientský software. Na trhu je k dispozici velké množství komerčních i open source řešení, a to velmi kvalitních. Popsat i jen většinu z těchto systémů by nebylo pro jejich velký počet možné, soustředil jsem se tedy alespoň na tři, z mého pohledu nejzajímavější, systémy.

OpenCms [GNU Lesser General Public License]

Tento systém je jedním z nejstaších moderních CMS systémů. Jeho původní verze byla dokončena v roce 1998 Alexanderem Kandziorem, který se následně v roce 2000 rozhodl projekt zveřejnit jako open source. OpenCms pak zaznamenalo velký rozvoj poté, co se do jeho vývoje zapojila řada společností i jednotlivců. V roce 2002 pak nad projektem převzala záštitu společnost Alkacon Software, založená původním autorem systému, Alexanderem Kandziorem.

Celý systém je postaven na technologii Java, jednotlivé stránky jsou řešeny pomocí JSP a šablon a obsah je pak uložen samostatně v souborech XML. Je možné jej provozovat jak v plně open source prostředích, jako jsou Linux, Apache, Tomcat či MySQL, tak i na komerčních řešeních (Windows NT, IIS, BEA Weblogic, Oracle DB). Kromě dvou zmíněných databázových serverů je možné použít i PostgreSQL nebo MSSQL, jako aplikační server poslouží i JBoss.

OpenCms umožňuje jednoduše vytvářet a spravovat komplexní webové stránky i bez jakékoliv znalosti html. Obsah může tvořit uživatel pomocí integrovaného WYSIWYG editoru, jehož prostředí se podobá standardním kancelářským aplikacím (typu Word a Excel), template engine se pak postará o správné zobrazování obsahu tak, aby zapadal do designu celého projektu. K samotnému obsahu pak uživatele přistupují pomocí webového prohlížeče odkudkoliv ze sítě, přičemž místa, ze kterých lze k systému přistupovat, je možné samozřejmě i přesně vyhranit.

Projekt je možno vyvíjet jako off-line, kdy k nim mají přístup pouze uživatelé, jež spoluvytvářejí jeho obsah, kteří tak mají jednoduchý přístup ke všem dokumentům v projektu, aniž by ten sám byl zveřejněn. Publikování se pak provede jednoduchým příkazem, jež automaticky všechny off-line dokumenty zveřejní, aniž by bylo třeba něco v systému kopírovat nebo nastavovat. Celý editovaný obsah zůstává na svém místě, pouze mu přibude příznak, jenž jej definuje jako zveřejněný. Měnit on-line a off-line příznak je samozřejmě možné i u jednotlivých dokumentů. U těch je také vedena historie všech změn a kdo je provedl, tudíž je možné i později libovolně přecházet mezi jejich staršími verzemi. V OpenCms je rovněž integrován workflow systém. Je možno zadávat úkoly jak jednotlivým uživatelům, tak i uživatelským skupinám, a celou historii řešení těchto úkolů sledovat.

Nedostatky OpenCms bych viděl hlavně v nepřiliš velké intuitivitě ovládání při vytváření obsahu a také nepřehlednosti a obtížnosti zdrojových kódů pro případ,

že by bylo třeba aplikaci pro vlastní potřebu něčím obohatit. I uživatelské prostředí v porovnání s novějšími systémy pro správu obsahu bledne a, ač jinak velice dobře funkčně vybaven, vzhledem k nim zaostává i v některých nabízených funkcích, podle mého názoru například nepříliš širokým modulem pro vyhledávání nad dokumenty. Všechny zmíněné problémy jsou však zřejmě způsobeny tím, že základ projektu je již takřka osm let starý.

[www.opencms.org]

Alfresco [Mozilla Public Licence]

Narozdíl od OpenCMS, Alfresco je velice mladým projektem. Vydán byl 31. října tohoto roku a sám sebe označuje jako revoluční na poli open source systémů pro správu obsahu. Předpoklady k tomu určitě má, společnost totiž byla založena odborníky, kteří se již na poli CMS dlouhou dobu pohybují, například spoluzakladatelem společnosti Documentum (v současnosti divize korporace EMC) Johnem Newtonem, či COO (chief operation officer) společnosti Business Objects Johnem Powellem. To je odlišuje od mnoha dalších podobných projektů, které jsou většinou vyvíjeny uskupeními na bázi komunit.

Projekt Alfresco je založen na nejnovějších Java technologiích. Použit je aplikační server Jboss 4.0, aplikační framework Spring 1.2, nástroj pro objektově relační mapování Hibernate 3.0, standardy JavaServer Faces / MyFaces, knihovna pro práci s dokumenty Apache Lucene 1.4 , která mimo jiné umožňuje jejich indexování a následné prohledávání či knihovna PDFBox, jež například umí získávat obsah z PDF souborů. Alfresco běží na všech databázových enginech, jenž jsou podporovány nástrojem Hibernate, včetně MySQL, Oracle či SQL Server. Celý systém je platformově nezávislý, je tedy možné jej používat jak ve Windows tak v operačních systémech Linux, MacOS a Unix, jako však ostatně většinu známějších CMS.

Alfresco určitě hned zaujme svým prostředím, které je velice moderně graficky zvtvárněné, přitom vypadá příjemně a přehledně. Systém používá rozhraní Microsoft CIFS (Common Internet File System), který umožňuje uživatelům pracovat s obsahem jako by pracovali se sdíleným diskem v prostředí Microsoft Windows. Narozdíl od ostatních rozhraní tak mohou uživatelé pracovat na obsahu off-line a ten po znovupřipojení synchronizovat. Tato technologie rovněž umožňuje, aby aplikace, které mají za normálních okolností přístup pouze k připojeným diskům či ne-síťovým file systémům, mohly s úložištěm systému Alfresco pracovat. Kromě CIFS je možné pro přístup k Alfrescu použít i jiná rozhraní, jako např. JSR-127 JSF, FTP, WebDAV či Web Services.

Co se týče funkcí, je na tom Alfresco také velmi dobře, podobně jako OpenCMS obsahuje podporu pro workflow a kontrolovatelné tvoření obsahu, má však i možnost jeho pokročilého prohledávání. Používá také technologii Smart Spaces (jak si ji tvůrci nazvali), tedy jakési "chytré prostory", jež mohou obsahovat soubory či další (pod)prostory. Těmto prostorům je možné stanovovat pravidla, jakým způsobem má být jejich obsah spravován (např. workflow, verzování nebo automatickou konverzi vložených dokumentů).

Systém Alfresco opravdu působí velice moderně a svěže, má přehledné webové stránky, jež nabízejí například velice užitečné flashové tutoriály pro představení nejvýznačnějších funkcí, ale i tutoriály pro vývojáře, kteří si plánují Alfresco upravit pro své potřeby, a uživatelskou příručku. Je rovněž možné si jeho funkce na požádání vyzkoušet v online demu. Dalším plusem je určitě použití nejnovějších open source technologií a velké množství užitečných funkcí včetně mnoha možných operací nad dokumenty, díky čemuž bývá Alfresco někdy řazeno i mezi DMS (Document Management Systems).

[www.alfresco.org]

Apache Lenya [Apache License 2.0]

Systém Lenya začal vyvíjet Michael Wechner na počátku roku 1999 pro účel správy obsahu vědeckého časopisu o vývoji vzorů (pattern formation). V roce 2000 Wechner spoluzaložil společnost Wyona, která pokračovala ve vývoji projektu Lenya pro internetovou verzi novin Neue Zürcher Zeitung. Na jaře roku 2003 věnovala Wyona Lenyu společnosti Apache Software Foundation, kde se z ní v září 2004 po nutném procesu „inkubace“, kterou musí podstoupit každý do Apache nově přijatý „člen“, stal vrcholný projekt. Co se týče názvu systému, je vytvořen kombinací jmen dvou synů původního tvůrce Michaela Wechnera, Leviho a Vanyaho.

Lenya je, podobně jako oba předešlé hodnocené produkty, postavena na Javě a standartech XML a XSLT, z dalších použitých technologií zmíním JSP a JavaScript. Jednou z komponent jádra systému je i framework Cocoon, který také spadá pod Apache Software Foundation. Java zajišťuje, že je možné Lenyu spustit na jakékoliv platformě, použít lze také širokou škálu webových i aplikačních serverů a databází.

Co se týče funkcí, tak podobně jako OpenCms obsahuje Lenya WYSIWYG editor pro jednoduché vytváření obsahu či kontrolu verzování včetně možnosti návratu ke kterékoliv dřívější verzi souboru. S projektem Alfresco má zase společné vyhledávání nad dokumenty pomocí Apache Lucene. Za velice zajímavý rys považuji možnost provozovat Lenyu přes více serverů, kdy např. jeden server slouží jako vývojový a jiný k vlastnímu publikování obsahu.

Pro projekt Lenya hovoří jasně to, že patří do „stáje“ Apache a může se tedy opírat o silnou podporu a rozsáhlou komunitu vývojářů i přispěvatelů. Nevýhodou je nepřiliš snadná instalace, kdy je třeba projekt buildovat, na rozdíl od dvou předešlých systémů, kdy se instalací jednoduše proklikáme pomocí průvodce. Pokud chceme místo přibaleného servletového kontejneru Jetty použít třeba Tomcat, je instalace ještě o něco složitější.

[lenya.apache.org]

Ostatní systémy

Výčet open source CMS systémů je opravdu velice rozsáhlý, možná také proto, že CMS je dosti široký pojem a mnohdy se mezi ně počítají např. i projekty DSpace či Fedora, které ale daleko spíše spadají na pole digitálních knihoven, a proto jsem v přehledu použil jiné významné systémy. Za zmínku by určitě stál i projekt MamboServer, jenž je, na rozdíl od tří recenzovaných produktů, postaven na technologiích PHP a MySQL. Velmi oblíbený je také robustní javovský InfoGlue, Magnolia či Daisy.

Ze tří mnou vybraných projektů na mě největší dojem udělal jednoznačně zbrusu nový přírůstek do rodiny CMS systémů, Alfresco. Když se vezmeme v úvahu, jak velký úspěch mělo vydání už jeho první verze, není jistě těžké uhadnout, že jej čeká zářná budoucnost.

Zdroje

- [1] CMS Matrix - www.cmsmatrix.org
- [2] Open Source software written in Java - www.roseindia.net/opensource
- [3] Enterprise Open Source - opensource.architectures.tv
- [4] The Free Encyclopedia - www.wikipedia.org
- [5] OpenCms Project - www.opencms.org
- [5] Alfresco - www.alfresco.org
- [6] Alfresco Software - www.alfresco.com
- [7] Apache Lenya - lenya.apache.org
- [8] InfoGlue Content Management Platform - www.infoglue.org
- [9] Mambo - www.mamboserver.com