ČESKÁ SHIBBOLETHOVSKÁ FEDERACE A VZDÁLENÝ PŘÍSTUP K ELEKTRONICKÝM ZDROJŮM POMOCÍ SHIBBOLETH

Jiří Pavlík a Petr Novák, ÚVT UK Praha

1. Úvod do problematiky vzdáleného přístupu

S rozšířením dostupnosti online připojení na Internet je běžné, že knihovny nabízejí také vzdálený přístup k nakupovaným databázím a dalším zdrojům. Základní způsob zpřístupnění – tedy ze studovny v knihovně – přestal brzy dostačovat. Uživatelé požadovali připojení ke zdrojům nakupovaným knihovnami z domova. Podobně, jako se proměňovala a rozvíjela nabídka producentů a poskytovatelů informačních zdrojů, měnily se i možnosti jejich vzdáleného zpřístupnění, měnila se i potenciální cílová skupina uživatelů – čtenářů, požadujících elektronické zdroje pomocí vzdáleného přístupu. Cílem knihoven je vzdálený přístup maximálně usnadnit a ulehčit jeho využívání pro situace, kdy není k dispozici okamžitá pomoc knihovníka specialisty. Zásadním požadavkem je pak bezpečnost poskytovaných zdrojů a umožnění přístupu pouze oprávněným registrovaným osobám tak, jak to definují sjednávané smlouvy a licence. Nabízíme lehký nástin minulých i aktuálních možností způsobů připojení s krátkým komentářem:

1.1. Terminálový přístup (telnet, ssh)

Určeno pro informační profesionály, nikoli běžné uživatele. Práce v textovém režimu, nepodporuje grafiku, nevýhodou je nutnost znalosti přesné syntaxe příkazů. Složité získávání plných textů – nutná navazující služba dodávání dokumentů. Omezení na databáze z aktuálně připojeného databázového centra. Výhody – rychlost, efektivita, práce s více databázemi najednou. Příklad – Dialog – webová verze http://www.dialogclassic.com/

1.2. Analogový modem – vytáčené spojení na stanice BBS

Komunita uživatelů Bulletin Board Stations – BBS provozovala veřejně dostupné systémy, k nimž bylo možno se připojit vytáčeným spojením. Počet aktuálně připojených uživatelů byl limitován počtem modemů na dané stanici BBS. Podporována byla základní textová ASCII grafika, modernější systémy umožňovaly připojování pomocí vlastních klientských programů. Bylo možno získávat i umisťovat plné texty a binární soubory

včetně programů. BBS byly postupně vytlačeny možnostmi Internetu.
Příklad – Infima – sloužila k distribuci dat Českého statistického úřadu.

1.3. Specializované klientské aplikace

Jednotlivá databázová centra postupně přecházela na připojení via Internet a protokoly TCP/IP. Pro přístup k informačním zdrojům byly využívány klientské aplikace, různé pro každého poskytovatele. Mezi jejich nevýhody patřily omezené množiny operačních systémů, bezpečnostní rizika vyplývající z uzavřeného kódu, nejasné specifikace protokolů, nemožnost návazných rozšíření. Výhodou byla možnost práce s více databázemi v grafickém prostředí, propojení bibliografických a plnotextových zdrojů, propojení online i lokálně instalovaných zdrojů, pohodlná práce s hesláři, tezaury, třídníky, nenáročnost na nastavení. Příklad: WinSpirs a MacSpirs, aplikace určené pro jednotný přístup původně k obsahu CD-ROM firmy SilverPlatter (odtud akronym SPIRS ze SilverPlatter Information Retrieval System). Dodnes používaná online varianta – WebSpirs.

1.4. Proxy servery a VPN

S rozvojem metody autentikace k informačnímu zdroji pomocí IP adresy bylo možno využít technologie proxy serveru, umožňujícího uživateli využívat IP adresu z rozsahu instituce s povoleným přístupem. Nutností byla konfigurace WWW prohlížeče či jiné aplikace.
VPN (Virtual Private Network) umožňuje tunelování, tedy provoz virtuálního síťového adaptéru instalovaného v systému a využívajícího připojení k Internetu. Provoz mezi virtuálním adaptérem a VPN serverem umístěným v cílové síti instituce je šifrován – tunelován. Toto řešení je robustní a komplexnější, přičemž vyžaduje znalosti protokolu TCP/IP či instalaci specifické klientské aplikace (dle druhu řešení, protokolu a způsobu autentizace).

1.5. Terminálový server

Tento způsob vyžaduje server, instalovaný v instituci, na kterém běží virtuální stroje s příslušnými aplikacemi. Pomocí klientské aplikace instalované na počítači uživatele je možné přistupovat k terminálovému serveru a spouštět jeho aplikace, přičemž dochází k přenosu grafického obrazu směrem k uživateli a vstupu klávesnice a myši směrem k serveru. Zvýšené požadavky na rychlost internetového připojení mohou být řešeny komprimací přenášeného obrazu. Variantou terminálového serveru je z operačního systému Windows známá aplikace Vzdálená plocha. Nevýhodou může být omezený počet přihlášených uživatelů.

1.6. Portálová řešení pro vzdálený přístup k elektronickým informačním zdrojům

Onelog – produkt britské společnosti Info Technology Supply Ltd. a Hidden Automatic Navigator (HAN) – produkt německé společnosti H+H Software GmbH nabízí podobnou množinu funkcionalit pro zpřístupnění elektronických informačních zdrojů. Obě aplikace nabízí širokou škálu nadstavbových funkcí často již integrovaných do stávajících knihovních systémů v institucích (správa licencí), případně vyžadují implementaci dalších komponent pro vzdálený přístup jak na straně instituce, tak na straně uživatele (Java, plug-in do prohlížeče, Citrix klient). Zároveň nepodporují Shibboleth jako standardní autentikační infrastrukturu.

1.7. Athens

Systém pro správu přístupu Athens[1] je provozován britskou agenturou EduServ, sdružuje více než 2000 akademických institucí a představuje defacto standard v této oblasti ve Spojeném království. Systém je podporován všemi velkými poskytovateli informačních zdrojů a jeho funkcionalita byla inspirací pro následný vznik single-sign-on projektů Shibboleth a Shibboleth2.

2 Shibboleth

2.1 Vznik a vývoj technologie, institucionální zabezpečení

Shibboleth je middleware vyvíjený organizací Internet 2 a nabízený jako freeware a open-source. Cílem Shibboleth je zajistit robutní, škálovatelnou, na standardech založenou infrastrukturu pro federativní autentikaci.

2.2 Princip technologie, WAYF/DS, IdP a SP

Principem Shibboleth je federativní autentikace, kdy uživatelé jsou při přístupu ke službám autentikováni v systémech domovských organizací. Na straně organizací, jejichž uživatelé využívají některé ze služeb, je implementována komponenta Identity Provider – IdP. Na straně služeb je implementována komponenta Service Provider – SP.

Organizace a služby jsou typicky zapojeny do tzv. federací. V rámci federace jsou evidovány IdP organizací a SP poskytovatelů služeb poskytovaných uživatelům z organizací zapojených do federace. Ve federaci jsou definována pravidla respektována zapojenými IdP a SP do federace. Pravidla či také politika federace určuje formát a význam informací poskytovaných IdP pro SP a zajišťuje také potřebný právní rámec.

[1] http://www.athensams.net/

Služba Where Are You From – WAYF, resp. Discovery Service – DS – nahrazující v aktuální verzi Shibboleth 2.0 službu WAYF z předchozích verzí, zajišťuje v rámci federace škálovatelné směrování uživatelů na IdP domovské organizace při požadavku přístupu k SP, k žádané službě.
V rámci shibbolethovských IdP je také typicky implementován některý ze Single-Sign-On systémů.

2.3 Aktuální stav vývoje v ČR, česká federace

V České republice je sdružením Cesnet provozována testovací federace czTestFed. Do provozu je připravována ostrá verze federace eduID.

3 EZproxy jako příklad knihovní aplikace využívající Shibboleth

3.1 EZproxy na Univerzitě Karlově

Univerzita Karlova předplácí 100 databází od cca 30 poskytovatelů při rozlišení 25 uživatelských skupin. Systém EZproxy byl vybrán, neboť nejlépe splňuje požadavky, kladené na systém zpřístupňující uživatelům elektronické informační zdroje.
Mezi požadavky patří:

snadná správa aplikace a její pokročilé funkce

- společná konfigurace zdrojů
- jednotná správa uživatelských přístupů integrovaná s celouniverzitními autentizačními službami a využití univerzitních účtů
- respektování uživatelských rolí (student, vyučující, zaměstnanec, uživatel knihovny, absolvent, student s přerušeným studiem aj.) v souladu s uzavřenými licenčními smlouvami
- odlišení přístupu k předpláceným zdrojům dle příslušnosti k fakultě či základní součásti (katedra nebo ústav) či jejich kombinace
- aplikační řešení funkční na vícero platformách (Linux, Windows 2003 server)
- nenáročnost na systémové zdroje
- rozšířená podpora nakládání s cookies a session ID
- příznivá cena (absence udržovacích poplatků)
- uživatelská komunita
- záznam uživatelských aktivit umožňuje vyhodnotit využívanost informačních zdrojů
- auditování a další postupy identifikující případné porušování podmínek užívání informačních zdrojů
- podpora neevropských znakových sad pro zdroje zpřístupňující texty v mimoevropských abecedách
- podpora zdrojů užívajících veškeré užívané techniky (JavaScript, Java Server Pages, Java applety, servlety, Flash, AJAX, OpenSearch, ASPX aj.)

- **požadavky na uživatele**
- jednoduché použití bez nutnosti složité konfigurace
- nezávislost na specifickém operačním systému, www prohlížeči, instalaci softwarových doplňků a komponent (pokud jejich instalaci nevyžadují cílové databázové aplikace a informační zdroje)
- komunikace na standardních TCP portech používaných pro http a https komunikaci

Integrace a podpora stávajících aplikačních řešení knihovních aplikací na UK

- Metalib UK
- SFX UK
- DigiTool UK
- ALEPH – CKIS UK
- 360 Search CERGE
- Google Scholar aj.

3.2 Jak EZproxy pracuje

EZproxy dynamicky připojuje k URL poskytovatelů databází svoji adresu, včetně zásahu do zdrojového kódu jednotlivých stahovaných nebo generovaných stránek. Díky konfiguraci lze nastavit rozsah modifikovaných (předkládaných) odkazů, například zdroje dostupné zdarma překládané být nemusí. Původní adresu http://www.zdroj.com server předloží uživateli jako http://www.zdroj.com.ezproxy.univerzita.cz. Uživatel následně pracuje s takto zpřístupněným zdrojem stejně, jako by pracoval na PC s IP adresou z rozsahu předplácející instituce. EZproxy nabízí také možnost zpřístupnění informačních zdrojů autentizovaných jménem a heslem. Pro tyto případy disponuje skriptem, který uživatele přihlásí vyplněním přihlašovacího formuláře.

3.3 Postup implementace

- Instalace operačního systému
- Zajištění konektivity a zálohování
- Konfigurace DNS serverů
- Instalace EZproxy
- Vyžádání a implementace testovací licence
- Vytvoření pokusných statických uživatelských účtů
- Nastavení testovacích zdrojů, ověření funkcionality
- Žádost o vystavení SSL certifikátů užívaných pro dva účely:
  - šifrování komunikace mezi centrálním autentizačním serverem (poskytovatelem identity) a EZproxy (poskytovatelem služby)
  - překlad URL zdrojů dostupných přes protokol https
- Import certifikátu do systému

- Oznámení implementace EZproxy serveru České Shibboleth federaci
- Nastavení federativních metadat
- Aktualizace metadat poskytovatele identity instituce
- Testování přihlašování
- Mapování uživatelských rolí z třídy eduPerson na uživatelské skupiny EZproxy
- Vytvoření uživatelských skupin elektronických informačních zdrojů
- Konfigurace informačních zdrojů autentizovaných přes IP adresu
- Konfigurace informačních zdrojů autentizovaných přes uživatelské jméno a heslo
- Konfigurace dalších knihovních aplikací instituce
- Testování a ladění
- Příprava školících a propagačních materiálů a nápovědy
- Školení
- Poskytování hotline pro uživatele EZproxy

3.4 Výsledek

EZproxy je na Univerzitě Karlově k dispozici pro nasazení v reálném provozu od akademického roku 2008–2009.

4 Závěr

EZproxy ve spojení se Shibboleth zajišťují pro uživatele pohodlné řešení autentikovaného a autorizovaného přístupu ke službám v prostředí WWW odkudkoli. Na straně uživatelů není potřeba instalace žádného dodatečného SW a podporovány jsou všechny běžné WWW prohlížeče v prostředí Mac OS X, Linux, MS Windows i dalších méně běžných systémech. Podpora pro Shibboleth je implementována či v přípravě u většiny významných poskytovatelů el. zdrojů, u systémů pro digitální repozitáře a u e-learningových systémů. Tam, kde Shibboleth ještě podporován není, Ezproxy zajistí transparentní autentikaci založenou na metodách IP adresa proxy serveru či institucionální jméno a heslo.

5 Zdroje

1. http://kis.is.cuni.cz/KSISENG-9.html
2. https://is.rice.edu/~bribbeck/Projects/NMI/NMI-EDITHAM-TMCCaseStudy.html
3. http://shibboleth.internet2.edu/
4. http://www.switch.ch/aai/
5. http://pez.cuni.cz
6. http://www.eduid.cz