



PROGRAM VEŘEJNÉ INFORMAČNÍ SLUŽBY KNIHOVEN

PODPROGRAM Č. 3

Rozvoj serveru obalkyknih.cz

Zpráva o plnění cílů projektu VÍSK pro rok 2013

Jihočeská vědecká knihovna v Českých Budějovicích

2014

Zhodnocení projektu

Projekt ObalkyKnih.cz sdružuje různé zdroje informací o dokumentech do jedné, snadno použitelné webové služby. Databáze aktuálně obsahuje přes **770 tisíc obálek a 75 tisíc obsahů** českých a zahraničních dokumentů. API služby využívá aktuálně přes 150 registrovaných knihoven a stovky neregistrovaných uživatelů.

Hlavním cílem projektu v roce 2013 bylo řešení extrémní zátěže na hardware a občasných výpadků dostupnosti. Viz příklad statistik využití služby:

Za říjen 2013 server vyřídil **52 057 553 dotazů** (poskytnutých obálek a obsahů), což odpovídá průměrně **19 dotazům každou vteřinu**. Největší zatížení serveru je v pracovní dny mezi 9:00 – 15:00, kdy průměrná hodnota odbavených dotazů za vteřinu dosahuje **40 požadavků**, ve špičkách dvojnásobek. Trvalý datový tok směrem k Internetu se pohybuje kolem **2 Mbps**.

V červnu 2013 bylo vypsáno výběrové řízení na dodavatele programátorských prací v rámci řešení rozvoje projektu Obalkyknih.cz, jehož předmětem bylo:

- Návrh architektury řešení nového serveru obalkyknih.cz umístěného v síti JVKCB. Cílem je zajištění dostatečného výkonu dotazovacího API a možnost škálovatelnosti do budoucna.
- Návrh nového datového modelu s orientací na využití NoSQL databází.
- Výběr a instalace konkrétní NoSQL databáze a aplikačního serveru pro funkční databázové a aplikační prostředí na serveru JVKCB.
- Zpracování funkčního a technického konceptu řešení nového serveru obalkyknih.cz.
- Realizace přechodu serveru obalkyknih.cz v JVKCB na 64-bitovou technologii.
- Realizace nové verze dotazovacího REST API rozhraní pro získávání obálek, obsahů, anotací, komentářů, hodnocení, případně dalších typů objektů dle konceptu, včetně podpůrné JavaScript knihovny pro zapojení do OPACu KIS Aleph.
- Realizace REST API rozhraní pro vkládání obálek dle konceptu na novém serveru obalkyknih.cz v JVKCB, ověřené vložení testovací obálky.
- Realizace REST API rozhraní pro vkládání obsahů, anotací, komentářů, hodnocení, případně dalších typů objektů dle konceptu na novém serveru obalkyknih.cz v JVKCB, ověřené vložení testovacích objektů.
- Zajištění synchronizace dat ze stávajícího serveru obalkyknih.cz umístěného v síti MZK, tzn. pravidelně aktualizovaná data na novém serveru obalkyknih.cz v JVKCB, získaná ze serveru v MZK.
- Optimalizace výkonu (rychlosti odezvy) pro vyšší zatížení doložená protokolem o zátěžovém testování.
- Pilotní implementace získávání obálek z nového serveru obalkyknih.cz do katalogu JVKCB jejímž výstupem bude funkční IPAC katalog JVKCB se správně zobrazovanými obálkami.
- Odevzdání zdrojových kódů a finální dokumentace projektu vložené do veřejného repozitáře na GitHub.

Do výběrového řízení se přihlásily 2 firmy a jako vítězná byla vybrána společnost Cosmotron Bohemia, s.r.o..

V létě 2013 došlo k přechodu z 32 na 64 bit platformu stávajícího serveru, což byl nezbytný předpoklad pro další postup. Následně byla zprovozněna kopie serveru a synchronizace mezi ostrým serverem běžícím v MZK a novým BackEndem v JVK.

Z důvodu extrémní zátěže serveru bylo nutné provádět vývoj a testy na kopii serveru.

Nové řešení aplikace obalkyknih.cz, nasazené do pilotního provozu nad ostrým katalogem JVKCB, bylo rozděleno na dvě části:

- ❑ **FrontEnd:** jen pro čtení, poskytuje data knihovním katalogům
- ❑ **BackEnd:** „zápisová část“ zajišťující import nových obálek z jednotlivých zdrojů, udržující hlavní datový model pro deduplikaci, provázání jednotlivých entit

Zároveň bylo nutné vyřešit vzájemnou synchronizaci FrontEndu a BackEndu.

Cílem bylo zajistit:

- dostatečný výkon dotazovacího API
- vybrat a implementovat technologii umožňující zajistit do budoucna odolnost x výpadkům
- možnost škálovatelnosti do budoucna
- přechod na 64-bitovou platformu
- jednodušší implementace do knihovních systémů
- zrychlení fungování a odezev webové služby

Z důvodu úspory finančních prostředků a vysokých požadavků na výkon byly pro realizaci softwarové stránky projektu vybrány open-source projekty **Node.js** a **mongoDB** běžících na 64bit OS **Debian**.

Posledním krokem bylo zprovoznění FrontEndu poskytujícím obálky a obsahy pro „ostrý“ knihovní katalog JVK a úprava dotazovacího API.

Popis řešení a veškeré kódy aplikace jsou **volně dostupné** jako open-source na adrese <https://github.com/cbvk/obalkyknih/wiki>. (Z důvodu přehlednosti nejsou součástí této zprávy).

Bylo pořízeno nové diskové pole pro provoz služby s důrazem na rychlost a bezpečnost - HP StoreVirtual 4330 - synchronní zápis na dvě samostatné jednotky v různých lokalitách. Celková kapacita dosahuje 2,4 TB (aktuálně projekt obsahuje cca. 1 TB dat, průměrný denní nárůst dat činí 1 GB). Data jsou zálohována na pořízený NAS store QNAP TS-879U-RP.

Zároveň došlo v JVK ke kompletní výměně centrální síťové infrastruktury (rychlost páteře navýšena na 10 GBit). Dohodou s poskytovatelem připojení k Internetu – společnost Cesnet – byla navýšena vnější konektivita ze 100 Mbit na 1000 Mbit. Náklady plně hradila JVK.

Cílem je zajištění dostatečného výkonu hardware pro přenos „ostrého“ serveru služby z MZK do JVK v průběhu roku 2014.

Pracovníci MZK vyvinuli v roce 2013 skenovacího klienta na obálky a obsahy a poskytli službu OCR prostřednictvím ABBYY Recognition Serveru MZK, čímž nahradili klíčovou část nefunkčního projektu TOC Národní knihovny a zároveň otevřeli možnost podílet se na skenování obálek a obsahů všem knihovnám.

V roce 2013 byla poskytnuta projektu řešeného v JVK dotace 578 000,-Kč na pořízení diskového pole, zálohovacího NASu a vývojářské práce (investiční prostředky). JVK z vlastního rozpočtu do projektu vložila 433 341,-Kč (42,85% podíl z celkových nákladů).

Detailní přehled nákladů projektu:

diskové pole HP StoreVirtual 4330	606 984,- Kč
zálohovací NAS QNAP TS-879U-RP	155 097,- Kč
programátorské práce	249 260,- Kč

1 011 341,- Kč

Další náklady projektu (hrazeny mimo projekt z rozpočtu JVK):

personální nároky (0,2 úvazku)	67 800,- Kč
server pro provoz, náklady provozu	177 905,- Kč
upgrade switchů, navýšení linky Cesnet (10% podíl)	82 999,- Kč

Celkem náklady	1 340 045 ,- Kč
Dotace MKČR – VISK 3	578 000,- Kč
Celkem náklady JVK v roce 2013	762 045,- Kč (57%)

Projekt byl prezentován na těchto knihovnických akcích:

11. září 2013 - Konference knihovny současnosti - Univerzita Palackého v Olomouci
21. listopadu 2013 - Setkání uživatelů ARL - Prušánky
26. listopadu 2013 - Porada ředitelů SDRUK - Hradec Králové
29. listopadu 2013 - Setkání uživatelů Souborného katalogu ČR - Městská knihovna v Praze

V Českých Budějovicích 13.1.2014

Ing. Jiří Nechvátal
Jihočeská vědecká knihovna v Českých Budějovicích