

Posudek diplomové práce

Diplomová práce *Radka Petrušky* s názvem *Použití Oracle RDBMS jako XML databáze* se zabývá návrhem vhodného úložiště XML dat v databázi Oracle. Cílem je otestovat na několika základních typech úložišť ukládání XML dat do databáze a vyhledávání v nich. Na základě získaných poznatků pak bylo úkolem diplomanta napsat modul pro informační systém IS/STAG, který bude do jeho databáze nahrávat data ze systému SIMS (Sdružené Informace Matrik Studentů), Na tomto importu ověří výsledky první části diplomové práce.

Diplomová práce je rozčleněna do čtyř částí. V první diplomant shrnuje teorii z oblasti XML. Zmiňuje zde všechny možnosti ukládání dat ve formátu XML. Jsou zde popsány tři základní možnosti ukládání XML dat do Oracle databáze (binární, nestrukturované a strukturované úložiště). Dále pak zde diplomat popisuje možnosti indexace XML dat a XML modely, které se používají pro zpracování XML dat v jazyce java.

V další části se diplomant zabývá realizací vlastního testování. Jsou zde uvedeny základní postupy použité při testování a také popis jednotlivých dílčích aplikací použitých při testování včetně vlastní finální aplikace – modulu pro IS/STAG.

V další části pak diplomant prezentuje své výsledky měření. Jsou zde porovnána binární, strukturované a nestrukturované úložiště a též další vlivy, které rychlost zpracování XML dat ovlivňují jako například složitost struktury XML, indexace, použité java datové struktury při operaci s XML dokumentem a také vliv počtu vláken při paralelním zpracování XML dokumentu. Dle očekávání nelze jednoznačně říci, který typ úložiště je nejlepší. Liší se to podle typu použitých XML dat.

Poslední část se věnuje reálnému testování naprogramovaného modulu pro IS/STAG. Zde byl kladen důraz na co nejrychlejší nahrání poměrně velkého XML souboru (stovky MB) do databáze Oracle. Dle závěrů diplomové práce se jeví jako nejvhodnější strukturované úložiště a to jak v oblasti vkládání dat, tak i v oblasti vyhledávání.

Teoretická část je v zásadě dobře zpracována, jsou zde vysvětleny všechny dále používané termíny a technologie. Diplomant zde mimo jiné velice dobře používá konkrétní vzorové SQL příkazy, které objasňují vlastnosti konkrétních objektů. Nicméně trochu nelogicky jsou dle mého názoru řazeny kapitoly na konci této části diplomové práce, které se již věnují vlastnímu testování, ale jsou ještě zařazeny do teoretické části práce. Ty bych spíš čekal v části následující. Vlastní testování bylo ale dle mého názoru dobře připraveno a výsledky lze považovat za relevantní. Osobně bych ale přivítal lepší prezentaci dosažených výsledků, např. formou srovnávacích grafů. Diplomová práce sice všechny dosažené výsledky obsahuje, ale občas je třeba hodně listovat a dohledávat srovnání. Přiložené programy fungovaly na první spuštění a obsahují i jednoduchou nápovědu. Naměřené časy odpovídají časům uvedeným v diplomové práci. Modul pro IS/STAG tudíž významně zrychlí nahrávání dat ze systému SIMS do IS/STAG. Po formální a grafické stránce nelze mít k diplomové práci žádné výraznější připomínky.

Diplomová práce *Radka Petrušky* s názvem *Použití Oracle RDBMS jako XML databáze* splňuje zadání v celém rozsahu a proto ji doporučuji k obhajobě a hodnotím ji stupněm

výborně

V Plzni 31. května 2016

Ing. Petr Jiroušek

SOUHLASÍ
S ORIGINÁLEM

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky