

Oponentní posudek diplomové práce

Autor práce: Ondřej Novák

Hodnocení:

Cílem diplomové práce je seznámení s problematikou vícevláknových a distribuovaných aplikací a následný návrh a implementace distribuovaného web crawleru. Velký důraz je kladen na robustnost a efektivnost software spolu s dodržováním všech pravidel kladených na web crawlery pro šetrný přístup k webovým serverům.

Předložená zpráva má kvalitní úpravu a až na pár drobných překlepů je gramaticky zcela v pořádku. V práci student názorně popisuje problematiku web crawleru a jeho distribuované varianty. Popis je srozumitelný a nechybí ani grafické znázornění problematiky.

Následně autor navrhuje architekturu distribuovaného web crawleru pomocí technologií FTP, SOAP, vícevláknového programování a síťového modelu klient/server. Zde bych očekával porovnání s dalšími síťovými architekturami a vysvětlení proč autor volí právě tuto architekturu. Nakonec autor uvádí omezující podmínky nutné pro funkčnost aplikace v reálném prostředí.

Čtvrtá kapitola obsahuje popis samotné implementace. Autor volí rozumnou míru detailu a názorně doplňuje popis funkčnosti grafickým znázorněním. Zároveň se nebojí zamýšlet nad odolností celého systému a možnými vlivy na výkonnost aplikace.

V poslední kapitole autor provádí testování implementace s pomocí dvou výpočetních strojů. Autor názorně demonstuje různé testovací scénáře, zejména zkoumá výkon aplikace při různých parametrech a také se snaží minimalizovat zátěž webových serverů v čase.

U obhajoby bych doporučil klást následující otázky:

- Jaké jsou síťové architektury pro distribuované zpracování a proč byla vybrána právě architektura klient/server?
- Jak autor řeší síťové výpadky klienta při procesu stahování URL z webu a při přístupu k serveru?

Až na uvedené drobné nedostatky bylo zadání splněno v celém rozsahu a práce byla v pořádku. Celkově doporučuji práci k obhajobě a hodnotím ji stupněm **výborně**.

V Praze dne 13.6.2011

Ing. Miroslav Holubec

