

Posudek oponenta diplomové práce

1. Identifikační údaje

Název práce: **Zpracování velkých objemů prostorových dat**

Autor práce: **Jakub Špidlen**

2. Cíl práce a jeho naplnění

Práce se zabývá problematikou zpracování velkých objemů prostorových dat. Cílem práce bylo navržení vhodného způsobu využití projektu Apache Hadoop pro zpracování velkých objemů prostorových dat a zároveň toto řešení prakticky ověřit.

Autor naplnil stanovené cíle práce a dodržel zásady pro zpracování diplomové práce.

3. Obsahové zpracování a přístup k řešení

Téma práce hodnotím jako nanejvýš aktuální. Vzhledem k tomu, že komerční sféra i veřejná správa shromažďuje obrovská množství dat, musí být nalezeny způsoby, jakými s těmito daty pracovat. Právě tímto problémem se zabývá předložená práce.

Z hlediska obsahu práci hodnotím jako nadprůměrnou. Zejména kapitoly 8 a 9 poukazují na autorovy odborné znalosti, kdy prokázal porozumění problematice distribuovaného zpracování dat. Práce není pouze popisná, autor nezůstal jen u návrhu systému, ale celý projekt realizoval v prostředí univerzitní laboratoře. Díky tomu je možno tuto práci považovat za ucelený návod pro řešení dané problematiky.

Práce je uspořádána logicky a svým rozsahem odpovídá řešenému problému. Autor zvolil správný přístup k řešení problému, přičemž využil odpovídající metody, postupy a nástroje. Autor se vyjádřil ke všem bodům zadání diplomové práce.

4. Formální náležitosti a úprava

Práce o rozsahu 76 stran textu je zpracována čistě, a srozumitelně. Lze jí vytknout minimum překlepů a nedostatků po stránce gramatické, stylistické a typografické. Ke snížené srozumitelnosti přispívá pouze nekonzistentní značení podkapitol viz. str. 14 a 28. Zde jsou nadpisy, které by velikostí písma odpovídaly druhé úrovni nadpisu, nejsou však číslovány.

5. Připomínky

Na straně 14 není zřejmé, proč autor popisuje geografické informační systémy v ČSOB pojišťovně, zda jsou data této pojišťovny specifická svým rozsahem a množstvím, nebo zda tato pojišťovna používá distribuované zpracování dat, nebo zda jde pouze o příklad využití geografických dat v komerční sféře.

Žádné další zásadní připomínky k práci nemám.

6. Hlavní otázky doporučené k bližšímu vysvětlení

- Na straně 19 hovoříte o tzv. spatial data cube (datové kostce pro práci s prostorovými daty), používá se tento model také ve spojení s distribuovaným výpočtem v Apache Hadoop? Znáte nějaký příklad?
- Existuje nějaká možnost pro uživatele ArcGIS, jak by mohli kombinovat analytické možnosti tohoto nástroje s využitím nástrojů pro zpracování velkoobjemových dat (např. Hadoop)?

Práci **doporučuji k obhajobě** a hodnotím ji stupněm

výborně

Pardubice, 24. května 2013

Ing. Hana Kopáčková, Ph.D.
Ústav systém. inženýrství a informatiky
Fakulta ekonomicko-správní
Univerzita Pardubice

