

POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Jméno studenta: Jaroslav Schnaubert

Název práce: Práce s velkými objemy dat v relačních a NoSQL databázích

Autor posudku: Hana Skalská

Cíl práce: 1. Popsat výhody relačních a nerelačních databází a použití pro různé typy úloh v praxi.

2. Zabývat se problémem optimalizace práce s velkým objemem dat, omezením a úskalím jednotlivých technologií a často řešenými problémy v reálném provozu.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

Nejsou zjištěné podobnosti (0 %)

Dílicí připomínky a náměty:

Práce vychází (str. 1) z rešerše tří podobně zaměřených prací. V úvodní části chybí metodika a východiska, která plynou pro další koncepci a strukturu práce. Nejsou vytýčené žádné mantinely pro východiska, která povedou k objasnění cíle 1 a dalších ze zadání práce. Vysvětlení má pravděpodobně plynout z kapitol 2 a 3, které podrobněji popisují vlastnosti dvou zvolených specifických zástupců každého typu (relační, nerelační) databází. Přehledné porovnání koncepčních rozdílů mezi oběma přístupy chybí.

Kapitoly 5 a 6 popisují vlastnosti a porovnávají MySQL a MongoDB pro velká data. Porovnání v kapitole 6 mohlo být v přehledné tabulce.

Až popis výsledků v závěru práce umožní porozumět struktuře a pojetí teoretické části.

Závěry práce mohly také přehledně shrnout výsledky porovnání a vytknout hlavní rozdíly.

Shrnutí výsledků mohlo zahrnout i poznatky z aplikace.

Seznam literatury by měl být řazený buď abecedně, nebo podle pořadí, ve kterém se citace poprvé použila, což zde není. Odkazy na firemní weby v zápatí pod čarou (např. strany 6 a 7) nejsou vhodné. V textu jsou drobné překlepy. Na koncích řádků se velmi často objevuje osamocená hláska, která by správně měla být svázaná pevnou mezerou s prvním výrazem následujícího řádku.

Celkové posouzení práce a zdůvodnění výsledné známky:

Práce obsahuje teoretickou a praktickou část. Koncepte práce a metodika nejsou zřetelně popsány, pro cíl 1 chybí specifikace kritérií porovnání, viz výše.

Cíl 2 je řešený pomocí konkrétních dvou úloh firemního prostředí. Specifikace dat a výsledky aplikace v kapitolách 5 až 7 popisují kroky, které vedly k jeho splnění. Je popsán postup manipulace s daty v omezeném časovém rozpětí nebo v živém provozu. V databázi MySQL jsou řešené úlohy vložení nového sloupce a hromadné úpravy hodnot (hledání a odstranění duplicit) v datové tabulce s rozsahem 500×10^6 řádků. V produkční databázi Mongo DB je řešená optimalizace indexů s cílem zrychlení načítání dat z několika kolekcí.

Závěrečné dvě části se zabývají porovnáním přístupů MySQL a MongoDB k optimalizaci dotazů a izolačních úrovní, následně monitorováním databáze a operací.

V porovnání mohlo být vysvětleno, zda některé rozdíly (které rozdíly) logicky plynou z koncepčních rozdílů mezi přístupy obou srovnávaných typů.

Práce splňuje většinu bodů ze zadání, popis dvou aplikací pro big data zvyšuje její hodnotu.

Otázky k obhajobě:

Použil jste pro aplikace vlastní vytvořené příkazy, nebo pracujete s typovými (například firemními) šablonami, které jste modifikoval?

Práci doporučuji k obhajobě.

Navržená výsledná známka: B

V Hradci Králové, dne 2. září 2019

podpis