

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Real time big data processing
Jméno autora:	Ilja Sheiko
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Ing. Jiří Šebek
Pracoviště oponenta práce:	Kabinet výuky informatiky (13142)

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Cílem této bakalářské práce bylo prostudovat principy ohledně dnes moderního tématu big data viz proudového zpracování událostí, zanalyzovat požadavky na systém, poskytující služby tohoto zpracování prostřednictvím user-friendly UI, vyhledat a prozkoumat open-source řešení, které tyto požadavky splňuje. Důležitou částí bylo otestovat dané řešení. Jedná se o praktickou práci zaměřenou na potřeby současných softwarových firem.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Na práci oceňuji, že student byl schopen v rámci jeho současné odbornosti rozumně zvážit všechny aspekty pro nasazení a využití projektu. Následně pak vybral technologie k projektu. Z obsahu textu je taktéž patrné, že student pochopil veškeré použité odborné texty a literaturu. Student zadání splnil, avšak explicitně řečené zadání v této práci chybí, ale to není chyba studenta, proto nezohledňuji na výsledném hodnocení.	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Úroveň bakalářské práce po odborné stránce odpovídá znalostem získaným nad rámec studia. Celá textová práce i s přílohami má 36 stránek.	
V textu mi někdy chybí citace při zadefinování některého termínu viz například termíny Big data, „Three V“ a jiné. (specificky by mělo být jasné z textu odkud se informace vzaly). Vyzdvihl bych velmi precizně sepsanou kapitolu 2, v které je mnoho informací a velmi se mi líbila (ať už zmíněné MapReduce, Kafka, Spark a jiné). Je vidět, že zde si student dal záležet.	
V části testování je zmíněné, že se testovalo na 3 uživatelích. Uživatelé byli vybráni z cílové skupiny. Není zde ale zmíněno proč je počet testujících 4. Jednalo se o kvalitativní nebo kvantitativní testování?	
V rámci testování byly provedeny i stress testy, což je pro dané téma důležité. Je škoda, že se v textu nevyskytují ukázky kódu nebo nastavení, které by byly v kontextu velmi zajímavé.	

Formální a jazyková úroveň, rozsah práce	B - velmi dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Typografická a jazyková úroveň je dobrá. Rozsah textu je dostatečný. Celá práce je na velice vysoké úrovni. Text je v anglickém jazyce a v některých částech chybí členy a/the a jiné drobnosti. Navzdory tomu se text dobře čte.	
Obrázek č. 1 je špatně čitelný, je potřeba zvětšit alespoň texty, aby bylo jasné co je na obrázku.	
V oficiálním zadání mi chybí text pro pokyny, ale to je spíše chyba vedoucího než studenta. Tabulky v textu nemají popisky	

viz tabulka v kapitole 5.4.2. Ve stejné kapitole je překlep v textu:

TP1 and TP2 were not stressful for the system, but TP3 results are much more interesting8. < - 8 nepatří do textu.

Výběr zdrojů, korektnost citací

C - dobře

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

U zdrojů převažují spíše internetové zdroje nad vědeckými články. Většina z nich je relevantních, ale citaci [20] obecně na wikipedii nepovažuji za relevantní. Wikipedie není dobrý zdroj informací a někdy ani tyto informace nemusí být pravdivé. Práce obsahuje 31 citací, což je dostačující. Literaturu 1-11 jsou správně ocitovány u ostatních bych přidal údaj kdy byla stránka s informací aktuální (může se stát, že stránka nebude už funkční apod.)

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Předložená bakalářská práce zpracovává téma velmi dobře, všechny aspekty jsou dobré, spíše nadprůměrné, avšak některé drobné problémy v textu jsou jak bylo popsáno výše. Tyto nedokonalosti ale nejsou zásadní. Zadání splněno bylo a student prokázal znalosti z oboru SI. Práce jako takové je velmi aktuální pro průmysl. Obecně bych hodnotil samotnou práci na A a textovou část na C proto dávám výslednou známku B.

Otázky k obhajobě:

- 1. V části testování je zmíněné, že se testovalo na 4 uživateli. Uživatelé byli vybráni z cílové skupiny. Není zde ale zmíněno proč je počet testujících 4. Jednalo se o kvalitativní nebo kvantitativní testování?*
- 2. Jaké jsou vize do budoucna tohoto projektu? Je možnost snadného rozšiřování?*

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 9.6.2019

Podpis: