

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Zpracování velkých dat v realtime režimu
Jméno autora:	Ilia Sheiko
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Vedoucí práce:	Ing. Marek Sušický
Pracoviště vedoucího práce:	Profinit EU s.r.o.

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce se zabývá rychle se rozvíjícím tématem realtime zpracování big dat. Existuje několik softwarových nástrojů, které dokáží se streamovanými daty pracovat a čím dál častěji se tyto technologie uplatňují v praxi. Cílem práce bylo provést rešerši existujících nástrojů, vybrat z nich nejvhodnější a následně provést experimenty. Vybraný nástroj pochází z Polska a odezvy od autorů nebyly ideální, dokumentace často neaktuální, či žádná, možnosti ladění omezené. Hlavně z tohoto důvodu hodnotím práci jako náročnější.	

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>	
Závěrečná práce splňuje zadání a bylo ověřeno, že výsledný systém je schopen zpracovávat velké množství informací zasílaných v reálném čase.	

Zvolený postup řešení	 vynikající
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Rešerše, jakým způsobem student dospěl k nástroji Nussknacker, je poměrně chudá – kapitola 4.1. Bylo by vhodnější uvést přehled streamovacích technologií s vhodnou licencí, tzn. Spark Streaming, Kafka streams a rozšířit hledání i na ně. Toto bylo v průběhu řešení provedeno, ale v písemné části práce to chybí. Z hlediska postupu řešení šlo o ukázkový SW projekt – na začátku proběhla analýza požadavků a několik schůzek s business zadavatelem. Na základě toho bylo navrženo řešení, které bylo průběžně předváděno, aby nedošlo k nedorozumění mezi zadavatelem a řešitelem. Součástí implementace byly i testy a performance benchmark, který potvrdil vhodnost a správnost řešení.	

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>	
Práce je implementační, student implementoval na interním clusteru firmy Profinit EU systém pro zpracování realtime dat s GUI rozhraním. V průběhu řešení se student seznámil s problematikou a je připraven řešit obdobné úlohy v reálném životě s reálnými daty, v reálných objemech u zákazníků.	

Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Jazykově obsahuje práce překlepy a gramatické chyby. Z hlediska typografické práce odpovídá požadavkům, byť například seznam požadavků je vhodné uvádět formou tabulky a nikoliv odrážkového seznamu. V příloze C jsou obrázky příliš malé a bylo by vhodnější je otočit, aby bylo možné přečíst popisky os. Rozsahem práce odpovídá požadavkům na bakalářskou práci, byť by písemná část mohla být o něco rozsáhlejší.	

Výběr zdrojů, korektnost citací

B - velmi 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.

Student po celou dobu řešení aktivně hledal a studoval materiály k dané problematice. Domnívám se, že 3 citace v práci jsou uvedeny správně a v souladu s citační etikou. Webové zdroje neobsahují datum, kdy došlo k získání dat.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Technické řešení je funkční a poskytuje výkon srovnatelný s komerčními systémy, které stojí miliony korun. Vše je navíc opensource a tudíž snadno rozšiřitelné a například využitelné i ve státní správě a místních samosprávách. Na práci je ta možné navázat implementací konektorů na některé systémy třetích stran a například SIEM systémy. Využití vidím v oblasti bezpečnosti, marketingu i financích.

Velice kladně hodnotím testy s uživateli, které jsou popsány v kapitole 5.3.2. Ze závěrů jsou vidět největší slabiny, které se pravděpodobně projeví v běžném provozu. V závěru autor říká, že je výsledek mnohem levnější, než čistá implementace, ale chybí odhad pracnosti Nussknackeru.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Student musel pochopit požadavky na CEP systémy, nastudovat technologie, kterými lze tyto požadavky splnit a následně realizovat benchmark. Vybraná technologie se zprvu zdála jako téměř funkční řešení a předpokládali jsme, že těžiště práce bude ve výkonovém ladění a performance měření. Bohužel se ukázalo, že informace na webových stránkách nebyly přesné a bylo nutné mnoho věcí dořešit a doimplementovat, než začal systém fungovat tak, jak bylo očekáváno. Vše ztěžovala dokumentace v Polštině a nekomunikativní autoři.

Následně byly implementovány některé funkční bloky počítající agregace nad zasílanými daty a provedeno měření. Dosažené výsledky jsou slibné natolik, že technologický stack Flink + Nussknacker budeme nabízet pro řešení CEP úloh.

Z hlediska reálného využití proto práci hodnotím velice kladně. Bohužel obsahuje gramatické chyby a některé další nedostatky, které byly shrnuty výše.

Otázky:

V závěru je hodnocení, že je výsledek mnohem levnější, než čistá implementace. Na kolik odhadujete pracnost Nussknackeru?

Jaké jsou nejvhodnější místa k dalšímu rozvoji tohoto systému?

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

Datum: 5.6.2019

Podpis: