

Big Data a NoSQL databáze

EDICE
PROFESIONÁL

 GRADA

**Irena Holubová, Jiří Kosek
Karel Minařík, David Novák**

Tuto knihu bychom rádi věnovali:

Kryštofovi.

– Irena

*Rodině, která mne podpořila při práci na knize,
i když dobře věděla, co ji čeká.*

– Jirka

Mým učitelům z Ústavu filosofie a religionistiky FF UK.

– Karel

Sofince, která mi rostla před očima spolu s knihou.

– David

Upozornění pro čtenáře a uživatele této knihy

Všechna práva vyhrazena. Žádná část této tištěné či elektronické knihy nesmí být reprodukována a šířena v papírové, elektronické či jiné podobě bez předchozího písemného souhlasu nakladatele.

**Doc. RNDr. Irena Holubová, Ph.D., Ing. Jiří Kosek, Mgr. Karel Minařík
a RNDr. David Novák, Ph.D.**

Big Data a NoSQL databáze

Knih je monografie

Vydala Grada Publishing, a.s.

U Průhonu 22, 170 00 Praha 7

tel.: +420 234 264 401, fax: +420 234 264 400

www.grada.cz

jako svou 6041. publikaci

Odborní recenzenti:

Doc. Ing. Michal Krátký, Ph.D.

RNDr. Jiří Materna, Ph.D.

Vydání knihy schválila Vědecká redakce nakladatelství Grada Publishing, a.s.

Odpovědný redaktor Petr Somogyi

Sazba Jiří Kosek

Grafické zpracování obrázků Milan Vokál

Návrh a zpracování obálky Vojtěch Kočí

Počet stran 288

První vydání, Praha 2015

Knih byla připravena v XML formátu DocBook

a vysázena pomocí XSL-FO v programu XEP.

Vytiskla Tiskárna v Ráji, s.r.o., Pardubice

© 2015 Grada Publishing, a.s.

Cover Photo © fotobanka allphoto

ISBN 978-80-247-5939-5 (ePub)

ISBN 978-80-247-5938-8 (pdf)

ISBN 978-80-247-5466-6 (print)

Stručný obsah

| | |
|------------------|----|
| 0 autorech | 13 |
| Předmluva | 15 |

I. Pojem Big Data a principy distribuovaného zpracování dat

| | |
|--|----|
| 1. Úvod | 19 |
| 2. Datové formáty | 29 |
| 3. Základní principy | 47 |
| 4. Zpracování dat pomocí MapReduce | 63 |

II. NoSQL databáze

| | |
|---|-----|
| 5. Základní principy NoSQL databází | 87 |
| 6. Databáze typu klíč-hodnota | 95 |
| 7. Dokumentové databáze | 109 |
| 8. Sloupcové databáze | 127 |
| 9. Grafové databáze | 143 |

III. Pokročilé aspekty zpracování Big Data

| | |
|--|-----|
| 10. Další aspekty zpracování Big Data | 171 |
| 11. Dotazování nad NoSQL databázemi | 193 |
| 12. Transakce v distribuovaném prostředí | 205 |
| 13. Pokročilé aspekty grafových databází | 217 |
| 14. Další databáze pro Big Data | 243 |
| Závěr | 261 |
| Použitá literatura | 263 |
| Rejstřík | 273 |

Obsah

| | |
|--|----|
| O autorech | 13 |
| Předmluva | 15 |
| I. Pojem Big Data a principy distribuovaného zpracování dat | |
| 1. Úvod | 19 |
| 1.1 Jak velká jsou Big Data? | 20 |
| 1.2 Historie a vznik NoSQL databází | 22 |
| 1.2.1 Konec relačních databází? | 25 |
| 1.3 O čem bude kniha | 27 |
| 2. Datové formáty | 29 |
| 2.1 JSON | 30 |
| 2.1.1 JSON schéma | 33 |
| 2.2 XML | 35 |
| 2.2.1 XML schémata | 37 |
| 2.3 YAML | 39 |
| 2.4 Formáty Linked Data | 41 |
| 2.4.1 RDF/XML | 41 |
| 2.4.2 JSON-LD | 42 |
| 2.5 CSV | 43 |
| 2.6 Optimalizace ukládání a přenosu dat | 44 |
| 2.6.1 Protocol Buffers | 44 |
| 2.6.2 Apache Thrift | 45 |
| 2.6.3 BSON | 45 |
| 2.6.4 EXI a FastInfoset | 45 |
| 2.6.5 ASN.1 | 46 |
| 2.7 Jaký formát vybrat | 46 |
| 3. Základní principy | 47 |
| 3.1 Škálovatelnost | 48 |
| 3.2 Konzistence | 49 |
| 3.2.1 Souběh transakcí | 50 |
| 3.2.2 CAP teorém | 52 |
| 3.2.3 Občasná konzistence | 54 |

| | | |
|-----------|--|-----------|
| 3.3 | Distribuce | 56 |
| 3.3.1 | Rozdělení dat (sharding) | 57 |
| 3.3.2 | Master-slave replikace | 58 |
| 3.3.3 | Peer-to-peer replikace | 59 |
| 3.3.4 | Replikace + sharding | 61 |
| 4. | Zpracování dat pomocí MapReduce | 63 |
| 4.1 | Funkce Map a Reduce | 66 |
| 4.1.1 | Další příklady | 68 |
| 4.2 | MapReduce framework | 68 |
| 4.2.1 | Další vlastnosti | 70 |
| 4.3 | Hadoop | 72 |
| 4.3.1 | HDFS | 72 |
| 4.3.2 | Hadoop MapReduce | 74 |
| 4.3.3 | Další nastavby systému Hadoop | 76 |
| 4.4 | Kritika a ústup od MapReduce | 82 |

II. NoSQL databáze

| | | |
|-----------|---|------------|
| 5. | Základní principy NoSQL databází | 87 |
| 5.1 | Společné principy NoSQL databází | 88 |
| 5.2 | Datové modely v NoSQL databázích | 89 |
| 5.3 | Typologie NoSQL databází | 93 |
| 6. | Databáze typu klíč-hodnota | 95 |
| 6.1 | Principy | 96 |
| 6.1.1 | Základní operace a práce s klíči | 96 |
| 6.1.2 | Jmenné prostory klíčů | 97 |
| 6.1.3 | Druhy úložišť typu klíč-hodnota | 98 |
| 6.2 | Realizace a vlastnosti | 99 |
| 6.2.1 | Distribuce dat | 99 |
| 6.2.2 | Konzistence a dostupnost dat | 102 |
| 6.2.3 | Lokální organizace dat | 105 |
| 6.3 | Práce s daty | 105 |
| 6.3.1 | Sekundární indexy | 106 |
| 6.3.2 | Redis | 106 |
| 7. | Dokumentové databáze | 109 |
| 7.1 | Datový model „dokument“ | 109 |
| 7.2 | Dotazování a manipulace s daty | 115 |
| 7.2.1 | Dotazy | 115 |
| 7.2.2 | Modifikace databáze | 116 |
| 7.2.3 | Agregované dotazy a MapReduce | 117 |
| 7.2.4 | Shrnutí | 117 |

| | | |
|-----------|---|------------|
| 7.3 | Vlastnosti dokumentových databází | 118 |
| 7.3.1 | Indexy | 118 |
| 7.3.2 | Replikace dat a dostupnost systému | 119 |
| 7.3.3 | Rozdělení dat | 122 |
| 7.3.4 | ACID pro jednotlivé operace a transakce | 123 |
| 7.4 | Závěr | 124 |
| 8. | Sloupcové databáze | 127 |
| 8.1 | Datový model | 128 |
| 8.2 | Cassandra: datový model sloupců v praxi | 132 |
| 8.2.1 | Data jako multidimenzionální pole | 133 |
| 8.2.2 | Data jako řídké tabulky | 134 |
| 8.3 | Struktura a vlastnosti systému | 136 |
| 8.3.1 | Distribuce a replikace dat | 136 |
| 8.3.2 | Lokální organizace dat | 137 |
| 8.4 | Dotazy, indexy a transakce | 138 |
| 8.4.1 | Dotazy | 139 |
| 8.4.2 | Indexy | 140 |
| 8.4.3 | Transakce | 141 |
| 9. | Grafové databáze | 143 |
| 9.1 | Typy grafů a související pojmy | 145 |
| 9.2 | Databáze Neo4j | 146 |
| 9.2.1 | Datový model Neo4j | 146 |
| 9.3 | Přístup k databázi Neo4j | 147 |
| 9.3.1 | Java API | 147 |
| 9.3.2 | Gremlin | 149 |
| 9.3.3 | Cypher | 152 |
| 9.4 | Pokročilé rysy Neo4j | 156 |
| 9.4.1 | Neo4j HA | 156 |
| 9.4.2 | Transakce | 157 |
| 9.4.3 | Indexy | 158 |
| 9.5 | Další grafové databáze | 162 |
| 9.5.1 | Sparksee | 163 |
| 9.5.2 | InfiniteGraph | 163 |
| 9.5.3 | OrientDB | 163 |
| 9.5.4 | Titan | 164 |
| 9.6 | RDF databáze | 164 |
| 9.7 | Srovnání úložišť pro grafy | 165 |
| 9.8 | Závěr | 167 |

III. Pokročilé aspekty zpracování Big Data

| | |
|---|------------|
| 10. Další aspekty zpracování Big Data | 171 |
| 10.1 Analytické zpracování Big Data | 172 |
| 10.1.1 Schéma dat | 172 |
| 10.1.2 Tvorba datových skladů | 175 |
| 10.1.3 Analytické zpracování | 176 |
| 10.2 Vizualizace Big Data | 178 |
| 10.2.1 Vizualizace propojených dat | 179 |
| 10.2.2 Nástroje pro vizualizaci | 181 |
| 10.3 Invertovaný index jako databáze | 182 |
| 10.3.1 Apache Lucene a jeho nástavby | 183 |
| 10.3.2 Zpracování logů | 186 |
| 10.4 Cloud computing | 187 |
| 10.4.1 Cloud computing a Big Data | 189 |
| 11. Dotazování nad NoSQL databázemi | 193 |
| 11.1 Přímý přístup pomocí programového rozhraní | 194 |
| 11.2 MapReduce | 196 |
| 11.3 Specifické dotazovací jazyky | 196 |
| 11.3.1 Elasticsearch Query DSL | 197 |
| 11.4 Univerzální dotazovací jazyky | 199 |
| 11.4.1 Deriváty SQL | 199 |
| 11.4.2 Rozšíření SQL | 199 |
| 11.4.3 XQuery | 202 |
| 11.4.4 JSONiq | 203 |
| 11.4.5 SPARQL | 203 |
| 11.5 Závěr | 204 |
| 12. Transakce v distribuovaném prostředí | 205 |
| 12.1 Vlastnosti CAP podrobněji | 205 |
| 12.2 Základní transakční modely | 206 |
| 12.2.1 Ploché transakce | 207 |
| 12.2.2 Zřetěžené transakce | 207 |
| 12.2.3 Hnízděné transakce | 207 |
| 12.3 Transakce v distribuovaném prostředí | 208 |
| 12.3.1 2PC protokol | 209 |
| 12.3.2 3PC protokol | 210 |
| 12.4 Optimistické a pesimistické off-line zámky | 210 |
| 12.4.1 Optimistický přístup | 211 |
| 12.4.2 Pesimistický přístup | 211 |
| 12.5 Uspořádání časových razítek | 213 |
| 12.5.1 Pesimistické uspořádání | 213 |
| 12.5.2 Optimistické uspořádání | 214 |

| | | |
|---------------------------------|---|------------|
| 12.6 | MVCC | 215 |
| 12.7 | Závěr | 216 |
| 13. | Pokročilé aspekty grafových databází | 217 |
| 13.1 | Reprezentace grafů | 217 |
| 13.1.1 | Matice sousednosti | 218 |
| 13.1.2 | Seznam sousedů | 218 |
| 13.1.3 | Matice incidence | 219 |
| 13.1.4 | Laplaceova matice | 219 |
| 13.2 | Lokalita dat | 220 |
| 13.3 | Distribuce grafu | 221 |
| 13.4 | Dotazování nad grafy | 223 |
| 13.4.1 | Typy dotazů | 224 |
| 13.4.2 | Vyhodnocování dotazů a indexace grafových dat | 225 |
| 13.4.3 | Dotazovací jazyky pro grafy | 234 |
| 13.5 | Závěr | 242 |
| 14. | Další databáze pro Big Data | 243 |
| 14.1 | Hybridní databáze | 243 |
| 14.1.1 | PostgreSQL | 244 |
| 14.1.2 | MarkLogic | 249 |
| 14.2 | Databáze ve webovém prohlížeči | 250 |
| 14.2.1 | Web Storage | 250 |
| 14.2.2 | Indexed Database | 252 |
| 14.3 | NewSQL databáze | 254 |
| 14.3.1 | VoltDB | 255 |
| 14.4 | Array databases | 256 |
| 14.4.1 | SciDB | 257 |
| Závěr | 261 | |
| Použitá literatura | 263 | |
| Rejstřík | 273 | |

O autorech

Doc. RNDr. Irena Holubová, Ph.D. se habilitovala v roce 2014 v oboru informatika na MFF UK v Praze, kde v současné době působí jako docent na Katedře softwarového inženýrství. Současně externě působí na Katedře počítačů FEL ČVUT. Je autorkou více než 80 původních článků, které byly publikovány na mezinárodních konferencích a v impaktovaných časopisech, z oblasti analýz reálných dat a operací, odvozování XML schémat, XML benchmarkingu, generování testovacích dat a efektivní propagace změn v komplexních systémech týkajících se převážně semi-strukturovaných dat. Čtyři z nich získaly významná mezinárodní ocenění. Je spoluautorkou knihy „Technologie XML“ vydané v roce 2008 v nakladatelství Grada, která v tomtéž roce získala Cenu děkana MFF UK za nejlepší monografii. V rámci svého pedagogického působení (spolu)vytvářela na MFF UK a FEL ČVUT předmět „Technologie XML“ a „Pokročilé aspekty a nové trendy v XML“. V nedávné době vytvořila nový předmět „Big Data management a NoSQL databáze“. Na tyto oblasti také zaměřuje vedené bakalářské, diplomové a dizertační práce. Více informací je možné nalézt na stránce <http://www.ksi.mff.cuni.cz/~holubova/>.

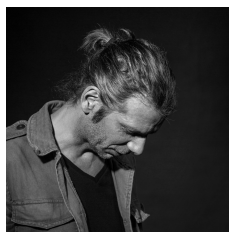


Ing. Jiří Kosek již více než 15 let poskytuje školení a konzultace v oblasti webových a XML technologií. Celá generace tvůrců webu vyrostla na jeho knížkách o HTML a PHP a je i autorem řady článků vydaných jak v Česku, tak v zahraničí. Na půdě Vysoké školy ekonomické v Praze vytvořil a učí předměty zaměřené na webové technologie a XML. Ve svém volném čase spolupořádá a programově zajišťuje konferenci XML Prague.¹ Jirka se podílí na tvorbě a údržbě důležitých standardů v několika organizacích – zejména W3C, OASIS a ISO a přispívá do několika open source projektů. Více se o jeho aktivitách můžete dozvědět na jeho stránkách <http://www.kosek.cz> a <http://xmlguru.cz>.



¹ <http://xmlprague.cz>

Mgr. Karel Minařík je webový designér a vývojář. Vystudoval filosofii na FF UK. Věnuje se programovacímu jazyku Ruby, využití nerelačních databází a vizualizaci dat. V současnosti pracuje pro společnost Elastic.² Žije v Praze. Více informací naleznete na webových stránkách <http://karmi.cz>.



RNDr. David Novák, Ph.D. získal doktorát z informatiky v roce 2008 na Fakultě informatiky MU v Brně, kde nyní pracuje jako vědecký pracovník. Ve svém výzkumu se věnuje zejména technikám pro podobnostní vyhledávání, vyhledávání v multimediálních datech a distribuovaným datovým strukturám. Pracoval na více než deseti národních a evropských výzkumných projektech a je autorem třiceti publikací na mezinárodních odborných fórech. V roce 2014 zavedl na FI MU předmět o NoSQL databázích. V roce 2015 získal Fulbrightovo stipendium na semestrální pobyt na University of Massachusetts Amherst v USA. Více informací naleznete na jeho stránce <http://disa.fi.muni.cz/david-novak/>.



² <http://elastic.co>

Předmluva

Kniha *Big Data a NoSQL databáze* má tři hlavní cíle: vysvětlit pojem Big Data, představit svět NoSQL databází a objasnit jeho souvislost s Big Data. Rozhodli jsme se ji napsat, protože zatím žádná ucelená publikace na dané téma v češtině neexistovala.

Knihu jsme rozdělili do tří částí. V první části vysvětlujeme fenomén Big Data a principy distribuovaného zpracování dat. Ve druhé představujeme několik typů databázových systémů označovaných jako NoSQL databáze. Poslední část knihy přináší přehled dalších nových typů databázových systémů, popisuje pokročilejší aspekty distribuovaného zpracování dat a nabízí přehled dalších souvisejících technologií.

Na první pohled různorodý autorský kolektiv spojuje právě dlouhodobý zájem o oblast zpracování dat. Irena Holubová a David Novák působí v akademickém prostředí. Big Data a související technologie vysvětlují v širších souvislostech. Přinášejí srovnání s tradičními technologiemi a ukazují, jak fenomén Big Data ovlivnil přístup ke zpracování dat a vývoj databázových systémů. Jiří Kosek se ve své praxi i v této knize věnuje zpracování strukturovaných dat, formátům pro jejich ukládání a možnostem dotazování. Karel Minařík sleduje vývoj NoSQL databází od jejich počátku a aktivně se ho účastní. Knihu obohatil zejména o cenné postřehy z praxe. Konkrétně se na obsahu knihy autoři podíleli následovně: Irena Holubová – kapitoly 1, 3, 4, 9, 12, 13 a sekce 10.2, 10.4, 14.3, 14.4; Jiří Kosek – kapitoly 2, 11 a sekce 14.1.2, 14.2; Karel Minařík – sekce 6.3.2, 10.3, 14.1.1 a většina šedých doplňujících rámečků; David Novák – kapitoly 5, 6, 7, 8 a sekce 10.1.

Na tomto místě bychom rádi poděkovali prof. RNDr. Jaroslavu Pokornému, CSc., doc. RNDr. Vlastislavu Dohnalovi, Ph.D. a RNDr. Martinu Svobodovi, Ph.D. za přečtení rukopisu textu a řadu cenných připomínek, které přispěly ke zkvalitnění výsledku. Dále děkujeme RNDr. Davidu Hokszozi, Ph.D., RNDr. Filipu Zavoralovi, Ph.D., RNDr. Jakubu Klímkovi, Ph.D., RNDr. Leu Galambošovi, Ph.D., Ing. Vladimíru Kyjonkovi, RNDr. Jakubu Lokočovi, Ph.D., Mgr. Jindřichu Mynarzovi a Lence Koskové Trískové za kontrolu vybraných kapitol textu a odborné konzultace k nim. V neposlední řadě pak patří velký dík recenzentům, doc. Ing. Michalu Krátkému, Ph.D. z Vysoké školy báňské – Technické univerzity Ostrava a RNDr. Jiřímu Maternovi, Ph.D. ze společnosti Seznam.cz. Děkujeme také sponzorům za významnou finanční pomoc. Autoři byli při přípravě knihy částečně financováni z Programu rozvoje vědních oblastí na Univerzitě Karlově (PRVOUK) č. 204-04/1204 (Irena Holubová).

Do knihy se promítají dlouholeté zkušenosti autorů z výuky kurzů zaměřených na zpracování dat, Big Data a NoSQL databáze. Látka byla Irenou Holubovou zpracována pro jednosemestrální kurz v inženýrské sekci MFF UK, který vznikl v roce 2012 a dnes je povinně volitelnou součástí magisterské výuky. Obdobný kurz vytvořil David Novák v roce 2014 na FI MU v Brně. V roce 2016 bude kurz v upravené podobě Irenou Holubovou vyučován i na FEL ČVUT. Studentům uvedených i obdobných kurzů bude kniha sloužit jako studijní opora.

Knihy obsahuje velké množství příkladů, které je možné nalézt i na webové stránce <http://www.ksi.mff.cuni.cz/bigdata>. Na stránkách naleznete i opravy případných chyb a další informace. Připomínky a dotazy ke knize můžete zasílat na adresu bigdata@ksi.mff.cuni.cz.

Přejeme vám příjemné čtení.

Autoři

Praha, Brno, Fryšava, Oldřichov v Hájích a Amherst, MA, USA, 25. září 2015

Část I.

**Pojem Big Data a principy
distribuovaného
zpracování dat**