

Distribuované spracovanie dát nad MapReduce architektúrou (Hadoop a Hive)

Martin Šeleng

Ústav Informatiky, Slovenská Akadémia Vied, Dúbravská cesta 9,
845 07 Bratislava, Slovakia
martin.seleng@savba.sk

Abstrakt. V prezentovanom príspevku v krátkosti predstavíme prostredie na paralelné a distribuované spracovanie dát založené na architektúre MapReduce (Hadoop implementácia) a distribuované skladisko Hive umožňujúce jednoduchšiu prácu nad architektúrou MapReduce (implementácia Hadoop) bez znalosti Javy. Projekt Hadoop bol inšpirovaný článkom spoločnosti Google, ktorá predstavili v roku 2004 vnútro svojho systému na spracovávanie odkazov na webových stránkach. Projekt Hive bol naštartovaný spoločnosťou Facebook na spracovanie veľkých objemov dát pomocou SQL jazyka.

Kľúčové slová: distribuované spracovanie dát, architektúra MapReduce, SQL, dátové skladisko

1 Úvod

Príspevok je rozdelený do dvoch častí, v ktorých postupne opíšeme distribuované skladisko Hive¹ a predvedieme praktické ukážky práce s týmto systémom. V článku nebudeme opisovať architektúru MapReduce, čitateľ môže nájsť viacej informácií v literatúre [1] a [4].

2 Distribuované skladisko dát Hive

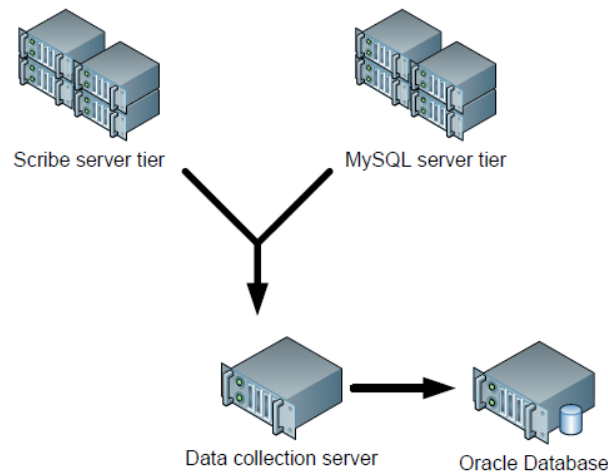
Projekt Hive bol naštartovaný spoločnosťou Facebook² (4 prispievatelia, všetci z Facebooku), z dôvodu spracovania veľkého množstva údajov v rámci sociálnej siete Facebook. Pôvodný proces spracovania dát spoločnosťou Facebook bol podľa nasledovný [3]:

- Dáta boli zbierané pomocou úloh zadaných v plánovači (boli to úlohy spúšťané v nočných časoch) a dáta boli zbierané do Oracle databázy.

¹ <http://hadoop.apache.org/hive/>

² <http://www.facebook.com/>

- ETL³ úlohy boli naimplementované v Pythone.
- V roku 2006 to bolo denne niekoľko 10 GB, v roku 2008 už 200GB nových dát, 5TB (komprimovaných) v roku 2009.



Obr. 1. Zbieranie používateľských dát zo sociálnej siete Facebook

V nasledujúcom odseku je uvedené na na čo všetko sa Hive dá použiť:

- spracovanie logov,
- dolovanie textových informácií,
- indexovanie dokumentov,
- modelovanie správania sa používateľov (túto vec má rád môj priateľ, možno ju budeš mať aj ty) a testovanie rôznych hypotéz.

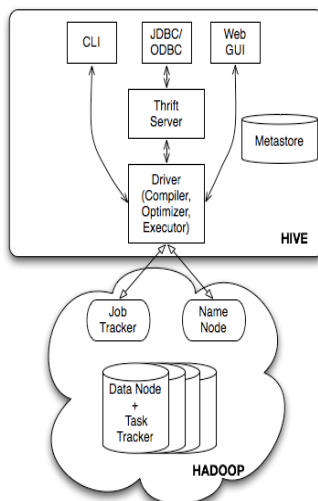
2.1 Hadoop ako distribuované dátové skladisko

V tejto časti si predstavíme ako sa dá Hadoop použiť ako distribuované skladisko pre dáta zo sociálnej siete Facebook. Dáta zo Scribe a MySQL servera sú nahrávané do HDFS a použijú sa MapReduce úlohy na spracovanie týchto dát. Čo v tomto návrhu chýba:

- jazyk, ktorým by sa dalo jednoducho tieto úlohy písať (bez potreby písať MapReduce programy),
- editor príkazového riadku, v ktorom by sa tieto úlohy mohli písať,
- schémy o jednotlivých tabuľkách v databázach

³ http://en.wikipedia.org/wiki/Extract,_transform,_load

Hive na všetky tieto otázky odpovedá (architektúru projektu Hive je možné vidieť na nasledujúcom obrázku).



Obr. 2. Architektúra systému Hive

Poskytuje vlastný editor príkazového riadku (tzv. hive>), ktorý je podobný MySQL editoru, ďalej poskytuje jazyk, ktorým je možné písať dopyty (zároveň je tu podpora aj pre JDBC klientov), uloženie metadát o databázach a tabuľkách. Najväčšou výhodou projektu Hive je možnosť písať SQL dopyty, pričom Hive preloží tieto dopyty do Map a Redcuc úloh (sú použité orientované acyklické grafy s veľmi jednoduchou optimalizáciou). Fyzicky je dátové skladisko uložené v HDFS v jednoduchých (plain) súboroch oddelených špeciálnym znakom alebo v špeciálnych sekvenčných súboroch⁴ v adresári: „/home/hive/warehouse“ a jednotlivé tabuľky sú uložené v podadresároch. Takisto je možné použiť vlastné formáty, pričom treba naimplementovať serializér a deserializér pre daný formát.

3 Ukážka práce so systémom Hive

V ukážke distribuovaného skladu dát použijeme diela zo stránky <http://zlatyfond.sme.sk>, pričom sme použili 2 diela od Pavla Országha-Hviezdoslava: Hájnikova žena⁵ a Ežo Vlkolinský⁶ (museli byť použité pdf verzie, pretože textové obsahujú množstvo chýb a preklepov).

⁴ <http://hadoop.apache.org/common/docs/current/api/org/apache/hadoop/io/SequenceFile.html>

⁵ http://zlatyfond.sme.sk/dielo/18/Hviezdoslav_Hajnikova-zena/1

⁶ http://zlatyfond.sme.sk/dielo/146/Orszagh-Hviezdoslav_Ezo-Vlkolinsky/1

V nasledujúcich krokoch si predstavíme, čo môžeme s týmito dvomi dielami, s Hadoop-om a dátovým skladosom Hive urobiť. Budeme používať MapReduce klaster s implementáciou Hadoop inštalovanou na UISAV. V čase písania článku sme mali k dispozícii 7 pracovných uzlov (slaves) a 1 riadiaci server (master).

- Predspracovanie dát ako napr.: konverzia pdf do textovej podoby, vyhodenie diakritiky a konverzia veľkých znakov na malé.
- Nahranie dát do HDFS.

```
$hadoop fs -copyFromLocal ezo.txt /user/hadoop/ezo.txt
$hadoop fs -copyFromLocal zena.txt user/hadoop/zena.txt
```

- Spočítanie slov (neberieme do úvahy slovenčinu, teda žiaden stemmer ani lematizátor) v jednotlivých dielach.

```
$hadoop jar hadoop-0.19.1-examples.jar grep ezo.txt ezo_freq '\w+'
$hadoop jar hadoop-0.19.1-examples.jar grep zena.txt zena_freq '\w+'
```

- Zmažeme logy, ktoré boli vytvorené Hadoop-om.

```
$hadoop fs -rmr ezo_freq/_logs
$hadoop fs -rmr zena_freq/_logs
```

- Prejdeme do editora príkazového riadku systému Hive a vytvoríme tabuľky zena a ezo do ktorých presunieme (LOAD DATA INPATH) dáta predspracované Hadoopom. Ak by sme chceli pôvodné dáta ponechať v HDFS tak použijeme príkaz: LOAD DATA LOCAL INPATH. Toto načítanie dát do tabuľky je veľmi rýchle (0,118 sekundy v našom prípade) pretože, sa jedná len o presunutie dát v rámci HDFS. V prípade že, sú dáta uložené v relačnej databáze, môžeme použiť ďalší z podprojektov projektu Hadoop Sqoop⁷ (pôvodne vyvíjaný spoločnosťou Cloudera⁸) na prevedenie relačných dát do HDFS.

```
$hive
hive> CREATE TABLE zena (freq INT, word STRING) ROW FORMAT DELIMITED
FIELDS TERMINATED BY '\t' STORED AS TEXTFILE;
hive> CREATE TABLE ezo (freq INT, word STRING) ROW FORMAT DELIMITED
FIELDS TERMINATED BY '\t' STORED AS TEXTFILE;
hive> LOAD DATA INPATH "zена_freq" INTO TABLE zena;
hive> LOAD DATA INPATH "ezo_freq" INTO TABLE ezo;
```

- Test, že dáta boli v poriadku načítané a takisto ukážka SQL dopytu v prostredí Hive. Tento SQL dopyt je rozdelený na dve MapReduce úlohy, ktorých vykonanie trvá na klastri 44,633 sekundy.

```
hive> SELECT * FROM ezo SORT BY freq DESC LIMIT 10;
```

⁷ <http://www.cloudera.com/downloads/sqoop/>

⁸ <http://www.cloudera.com/>

Všetky znaky malé		Aj veľké znaky	
Slovo	Počet	Slovo	Počet
sa	656	sa	656
a	598	a	440
i	427	v	362
v	375	i	352
na	322	na	303
co	271	si	242
to	254	co	232
si	245	to	214
len	201	s	182
tak	195	len	172

- Najčastejšie sa vyskytujúce frekvencie slov v diele Hájnikova žena. Tento dopyt je takisto rozdelený na 2 úlohy a jeho vykonanie trvá 45,718 sekundy.

```
hive> SELECT freq, COUNT(1) AS f2 FROM zena GROUP BY freq SORT BY f2
DESC;
```

Všetky znaky malé		Aj veľké znaky	
Počet slov	Počet výskytov	Počet slov	Počet výskytov
8727	1	9153	1
1805	2	1854	2
691	3	704	3
326	4	355	4
190	5	191	5
137	6	137	6
83	7	94	7
75	8	66	8
38	11	38	11
38	9	37	9
38	10	36	10

- Orientovaný acyklický graf vykonávania „SQL“ dopytu získame nasledovným príkazom

```
hive> EXPLAIN SELECT freq, COUNT(1) AS f2 FROM zena GROUP BY freq
SORT BY f2 DESC;
```

- Ďalšou vecou, ktorú Hive dokáže je ďalší „SQL“ príkaz JOIN. Najprv vytvoríme spoločnú tabuľku a následne obidve spojíme do jednej (samozrejme slová musia byť rovnaké). Toto spojenie je jedna MapReduce úloha, ktorá trvá 22,451 sekúnd.

```
hive> CREATE TABLE spojenja (word STRING, ezo_f INT, zena_f INT);
hive> INSERT OVERWRITE TABLE spojenja SELECT e.word, e.freq, z.freq
FROM ezo e JOIN zena z ON (e.word = z.word);
```

- Posledná vec bude zistenie, ktoré slovo sa najčastejšie vyskytuje v obidvoch dielach spolu, pričom vypíšeme aj jednotlivé frekvencie. Tento dopyt je takisto rozdelený na 2 úlohy a jeho vykonanie trvá 44,697 sekundy.

```
hive> SELECT word, ezo_f, zena_f, (ezo_f + zena_f) AS s FROM spojenja
SORT BY s DESC LIMIT 10;
```

Všetky znaky malé				Aj veľké znaky			
Slovo	Počet slov Ežo	Počet slov Žena	Počet slov spolu	Slovo	Počet slov Ežo	Počet slov Žena	Počet slov spolu
sa	656	970	1626	sa	656	970	1626
a	598	865	1463	a	440	742	1182
v	375	809	1184	v	362	786	1148
i	427	521	948	na	303	499	802
na	322	535	857	i	352	420	772
co	271	347	618	co	232	281	513
jak	118	432	550	jak	113	381	494
to	254	248	502	to	214	214	428
len	201	242	443	si	242	166	408
tak	195	238	434	z	158	249	407

4 Záver

V príspevku sme predstavili dátové skladisko postavené nad architektúrou MapReduce a ukázali ako sa relatívne jednoducho a distribuovane dá s dátami v nich uloženými pracovať. Okrem systému Hive existujú aj iné, ktoré poskytujú rovnakú alebo podobnú funkcionálnosť ako sú napr.: Pig⁹ (podprojekt projektu Hadoop vyvíjaný hlavne spoločnosťou Yahoo!¹⁰), Sawzall¹¹ od spoločnosti Google, JAQL¹² od spoločnosti IBM, Scope¹³ od spoločnosti Microsoft a YAML MapReduce od firmy Greenplum¹⁴.

Podakovanie: Táto práca je podporovaná projektmi: RECLER ITMS: 26240220029, SMART ITMS: 26240120005, SMART II ITMS: 26240120029.

- [1] Dean, J., Ghemawat, S.: MapReduce: Simplified Data Processing on Large Clusters, <http://labs.google.com/papers/mapreduce.html> (Prístup Sept. 2010)
- [2] Ghemawat, S., Gobioff, H., Leung, S-T.: The Google File System <http://labs.google.com/papers/gfs.html> (Prístup Sept. 2010)
- [3] <http://www.cloudera.com/wp-content/uploads/2010/01/6-IntroToHive.pdf> (Prístup Sept. 2010)
- [4] Šeleng, M., Laclavík, M., Hluchý, L.: Použitie MapReduce architektúry na spracovanie veľkých informačných zdrojov. In WIKT 2008: 3rd Workshop on Intelligent and Knowledge Oriented Technologies. Editor Pavol Návrat, Valentino Vranič. - Bratislava : Nakladateľstvo STU Bratislava, 2009, p. 27-34. ISBN 978-80-227-3027-3

⁹ <http://hadoop.apache.org/pig/>

¹⁰ <http://research.yahoo.com/node/90>

¹¹ <http://research.google.com/archive/sawzall.html>

¹² <http://www.almaden.ibm.com/cs/projects/jaql/>

¹³ <http://www.cs.uwaterloo.ca/~kmsalem/courses/CS848W10/presentations/Aluc-Scope.pdf>

¹⁴ <http://www.greenplum.com/technology/>