

Functional Programing in Concurrent Distributed Systems

Final Project Report File

פרויקט 7

עיבוד מקבילי ומבוזר של Big Data מ-IMDB באמצעות אלגוריתם MapReduce

עומר לוכסמבורג 205500390

עילי נוריאל 312538580

תיאור המערכת

מטרת הפרויקט

שליטה בשפת ארלנג, עבודה בצורה מקבילית ומבוזרת וכן עבודה עם מידע מבוזר. בנוסף, מטרת הפרויקט היא להבין את אלגוריתם ה-Map-Reduce, שימומש על נתונים מתוך אתר IMDb. רעיון המערכת הוא שכל לקוח (client) יוכל להרכיב שאלה כללית אודות נתונים הרלוונטיים לסרטים מהאתר ולקבל תוצאה רלוונטית לפי הפרמטרים המבוקשים.

רעיון כללי

מערכת IMDb Map-Reduce מחולקת לשלושה חלקים – מאסטר, שרתים, ולקוחות, כאשר המאסטר הוא המחבר בין כולם.

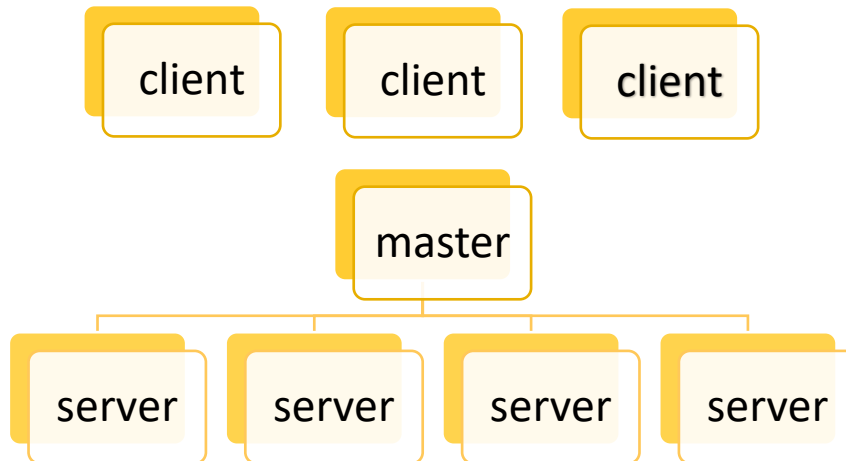
המאסטר מנהל את התקשורת מול השרתים ומול הלקוחות. תפקידו הוא לשמש כמוניטור על עליה ונפילה של שרתים – ובהתאם לכך שליחת המידע לביזור מחדש, וכן מענה ללקוחות בזמן אמת, באופן מקבילי, על ידי שליחת הבקשה לכל השרתים המחוברים למערכת.

השרתים מקבלים את המידע הרלוונטי אליהם ושומרים אותו במחשב המריץ את התהליך. כאשר תגיע שאילתה מהמאסטר, יוקצה תהליך שיממש אלגוריתם Map-Reduce רלוונטי על המידע המקומי. את התוצאה השרת ישלח למאסטר והמאסטר יאחד את המידע ויחזיר אותו ישירות ללקוח.

הלקוח, שולח את הבקשה ומקבל מענה. בנוסף למענה היבש לפי הבקשה ששלח, יתווספו עוד סטטיסטיקות רלוונטיות לתצוגת הלקוח.

מבנה המערכת

מטרת הפרויקט הראשונית הייתה לפתח את המערכת הנ"ל בעזרת מערכת ה-Disco¹. מאחר והמערכת לא פעלה, מימשנו מערכת דומה ואיכותית בשפת ארלנג בלבד המושתת על אותם קווי ייסוד:



כל לקוח מתקשר עם המאסטר, והמענה לו יתבצע באופן נפרד ומקבילי משאר הלקוחות. כך יתבצעו גם הבקשות שמתקבלות בשרתים, כדי לא ליצור צוואר בקבוק של מענה בודד ללקוח מתוך המאסטר.

המערכת פותחה בשפת ארלנג (Erlang 22, OTP22, erts-10.7).

1. **מאסטר** – סינגלטון – מריץ את קובץ `master.erl`, ואת `dataDistributor.erl`, `parse_csv.erl`. ממומש ב-`gen_server`.

- אחראי על התקשורת בינו לבין הלקוחות בזמן אמת – מקבל שאילתה, מחזיר את המגולם.
- אחראי על התקשורת בינו לבין השרתים בזמן אמת – מעביר פקודות לביצוע, מקבל תשובה מכל אחד.
- מנהל את ביזור המידע הדינאמי – מתוך רשימת `serverslist.txt` יזהה אילו שרתים מחוברים. אל שרתים אלו ישלח המידע. כאשר שרת מתוך הרשימה עולה, יבוצע ביזור מחדש של המידע באופן שווה. כאשר שרת המחזיק מידע נופל, המידע יבוזר מחדש כדי לא לאבד אף מידע המוחזר ללקוח. כל הביצוע הנ"ל קורה באופן אוטומטי ע"י קובץ `input` ידני `serverslist.txt` בלבד.

טיפול בהודעות – צד המאסטר

המאסטר מקבל הודעת `query` מהלקוח, ומעביר אותה הלאה. בנוסף המאסטר מקבל לתהליך שנוצר עבור שרת ספציפי את ההודעה `Reply/table_error` – הודעה זו היא התשובה של השרת

¹ מתוך: <http://discoproject.org>

לשאלתה שהמאסטר שלח, את התשובה הזאת נשלח עם כל ההודעות מכלל השרתים ישירות ללקוח.

בנוסף המאסטר מקבל הודעות `nodeup`, `nodedown` הרלוונטיות רק ל-`node`-ים בשם `server#`. אם התקבלה הודעה כזו, המאסטר מבצע בהתאם ביזור מחדש של המידע לצורך שמירה על איזון בין המידע ומניעת איבוד מידע שאוחסן בשרת שנפל.

במידה ונפל שרת בזמן עבודה על שאילתה, כלומר בזמן שהמאסטר מחכה לתשובה ממנו, המאסטר יקבל הודעת `nodedown` בתהליך ייעודי לניטור שרת זה – שיעביר הודעה לתהליך הראשי כי אין עליו להמשיך לחכות לתשובה משרת זה – והמאסטר יחזיר ללקוח את התוצאות משאר השרתים.

2. שרת (server) – מרובה – מריץ את קובץ `server.erl` ואת `mapreduce.erl`. ממומש ב-`gen_server`.

- אחראי על שמירת המידע הלוקלי במחשב המארח.
- אחראי על פיענוח שאילתה, וביצוע אלגוריתם Map-Reduce בהתאם לדרישה.
- אחראי על שליחת המידע חזרה בצורה קומפקטית ככל הניתן, אל המאסטר.
- שולח הודעות `nodeup`, `nodedown` למאסטר, לצורך ביזור דינאמי של המידע (כפי שפורט לעיל).

טיפול בהודעות – צד שרת

שרת מקבל את ההודעה `store` – הודעה זו נשלחת מהמאסטר לצורך שמירת המידע שבהודעה בתור מידע לוקלי במחשב המארח. לאחר שמירה זו השרת מוכן לקבל הודעות שאילתה לצורך ביצוע Map-Reduce.

השרת מקבל בנוסף הודעות `query` – הודעות מסוג זה שולחות שאילתה לביצוע. השרת פותח תהליך חדש המבצע את אלגוריתם ה-Map-Reduce הרלוונטי לפי השאילתה, על המידע המאוחסן במחשב המארח, ומחזיר תשובה ישר כאשר היא התקבלה.

3. לקוח (client) – מרובה – מריץ את קובץ `wxclient.erl`. ממומש בממשק GUI ע"י `wxwidgets`.

- ניהול התצוגה הגרפית והממשק למשתמש.
- אחראי על ניתוח השאילתה ושליחתה למאסטר.
- את המידע מציג באופן גנרי בחלון חדש, לפי הנתונים שהוחזרו.
- מציג עם התוצאות סטטיסטיקות נוספות אודות השאילתה.

טיפול בהודעות – צד לקוח

ללקוח הודעה הנשלח למאסטר – הודעת `query`. בעקבות הודעה זו נקבל את המידע הרלוונטי אל הלקוח, מהמאסטר. שאילתה זו מכילה מידע אודות הקטגוריה בה אנו מחפשים והערך שהמשתמש הכניס (`string`), בנוסף לכך היא מכילה את הקטגוריות הרצויות להצגה. לפני שליחת הבקשה, נוכל לקבוע את אופי מיון התוצאות כרצוננו.

במידה והמאסטר נפל טרם שליחת השאילתה, או במהלכה – הקליינט יקבל הודעת nodedown בתהליך ייעודי לניטור המאסטר – שיעביר הודעה לתהליך הראשי כי אין עליו להמשיך לחכות לתשובה וכי המאסטר לא זמין.

רצף התוכנית

כאשר המערכת עולה (הכוונה למאסטר) המאסטר מבזר את המידע ההתחלתי המוצג כקובץ CSV לפי המחשבים הרשומים לו ברשימת *serverslists.txt*, אשר מחוברים לרשת. לאחר גילוי הסרברים, המאסטר מבזר את המידע (לאחר shuffle - כדי למנוע עומס על סרבר מסוים בלבד) ושולח לכל סרבר chunk-ים של מידע בגודל זהה. מאותה עת, כל פעם שיעלה או ייפול node של שרת, המערכת תבצע ביזור מחדש של המידע לצורך איזון ושלמות המידע.

כעת ניתן להעלות ולהוריד שרתים באופן נטול חריגות. לצורך חיבור לקוח, נריץ את ממשק ה-GUI שרשום ב-wxclient.erl. מממשק זה, ניתן לבצע שאילתות ישירות למאסטר. את השאילתות המאסטר יעביר לכל השרתים הזמינים, וימתין להודעה. בזמן ההמתנה להודעה, המאסטר ממשיך לקבל בקשות באופן בלתי תלוי (לפי יכולת המעבד המריץ אותו).

כל שאילתה תענה ללקוח בחלון חדש המציג את כל הנתונים הרצויים. סגירת חלון הממשק, תהווה ניתוק של הלקוח באופן אוטומטי מהמאסטר.

ממשק המשתמש

ממשק המשתמש בנוי משני חלונות.

הראשון הוא חלון השאילתה (Query Window) בו המשתמש ממלא שאילתות הנשלחות אל המאסטר. חלון זה בעל מנגנון תקינות הקלט, על מנת שחשובים שיעשו בהמשך האלגוריתם לא יחזירו חריגה. החלון השני הוא חלון התוצאות. בחלון זה מרוכז כל המידע אודות הסרטים העונים על השאילתה שנשלחה. בנוסף לכך נתונים עוד סטטיסטיקות שונות בנוגע למידע המוחזר.

1. חלון בקשת השאילתה:

The screenshot shows the 'IMDb Map-Reduce Project' window with a 'Query Window' tab. It contains several input sections and a 'Send Query' button. Yellow arrows point from specific UI elements to explanatory text boxes on the right.

- Insert Value (case-sensitive):** A text input field. An arrow points to a box containing: הכנסת הערך לפיו מחפשים מידע.
- Title, Year, Genre, Duration:** A list of categories. An arrow points to a box containing: הקטגוריה של הערך אותו מחפשים.
- Select Categories (ID, Title are always selected):** A list of checkboxes for various categories. An arrow points to a box containing: בחירת המידע שיוצג בחלון התוצאות על כל הסרטים התואמים לחיפוש.
- Select Sorting Category:** A list of categories for sorting. An arrow points to a box containing: בחירת קטגוריה לפיה ימוינו הרשומות בחלון התוצאות.
- Send Query:** A button at the bottom. An arrow points to a box containing: שליחת השאילתה.

2. חלון התוצאות

IMDb Map-Reduce Project

Results Window

Search value: Drama | Value category: Genre | Sorting by: Year

Result's Statistics:
 47110 Results | 3 Servers | Evaluation Time: 3241ms | Organize Time: 117543
 Year - Min 1906 | Max 2020 | Avg 1993.55
 Duration - Min 45 | Max 808 | Avg 103.77
 Score - Min 1.0 | Max 9.8 | Avg 6.24

	ID	Title	in (m)	Country	Actors	Score	Year
1	tt0000574	The Story of the	70	Australia	Elizabeth T	6.1	1906
2	tt0001892	Den sorte dr	53	Germany, D	Asta Nielse	5.8	1911
3	tt0002130	L'Inferno	68	Italy	Salvatore P	7.0	1911
4	tt0266688	Karadjordje	80	Serbia	Jovan Anto	6.2	1911
5	tt0002199	From the Mang	60	USA	R. Hendersc	5.7	1912
6	tt0319904	Le mystère de	45	France	Suzanne Gr	6.8	1912
7	tt0002101	Cleopatra	100	USA	Helen Gard	5.2	1912
8	tt0002461	Richard III	55	France, USA	Robert Gen	5.5	1912
9	tt0003471	Traffic in Souls	88	USA	Jane Gail, E	6.0	1913
10	tt0368309	Sumerki zhensk	48	Russia	Nina Chern	6.6	1913
11	tt0002844	Fantomas - À	54	France	Renée Nav	7.0	1913
12	tt0002646	Atlantis	121	Denmark	Olaf Fjell	6.6	1913
13	tt0003489	Gli ultimi giorni	88	Italy	Fernanda N	6.2	1913
14	tt0003014	Il calvario di una	96	Sweden	Hilda Borgs	7.1	1913
15	tt0003883	L'enfant de Paris	124	France	Léonce P	7.5	1913
16	tt0003165	Le mort qui tue	90	France	Renée Nav	7.0	1913
17	tt0003419	Lo studente di	85	Germany	Paul Weger	6.5	1913
18	tt0002445	Quo Vadis?	120	Italy	Amleto No	6.2	1913
19	tt0002027	Juve contre Fan	61	France	Renée Nav	7.0	1913

נתונים על השאילתה, שהוכנסה על ידי המשתמש

נתונים סטטיסטיים כלליים על ביצוע השאילתה וכן נתונים נוספים על תוצאות החיפוש

טבלה המציגה את תוצאות החיפוש ממוינים לפי השדה שנבחר (מהקטן לגדול) – כאן ניתן לראות שהמידע מסודר לפי Year.

- Evaluation Time – הזמן מרגע שליחת השאילתה, עד לקבלת כל הנתונים בלקוח.
- Organize Time – הזמן מרגע קבלת כל הנתונים, עד להצגתם במסך המשתמש (כולל מיון הנתונים, ניתוח סטטיסטיקות ותצוגה ב-GUI).

ראינו כי זמן האבליואציה לוקח בערך אותו הזמן עבור מחשבים שונים באותה תצורת מערכת. לעומת זאת, זמן הארגון של המידע משתנה ממחשב למחשב באופן דרסטי יותר.

מבני נתונים

אנו משתמשים בשמירת המידע ע"י שמירת טבלת ets לקובץ. שימוש זה מאפשר לנו לשלוח ערכים כמשתני ארלנג שהוגדרו בביזור המידע באופן יעיל ומהיר. כל chunk של טבלה נשמר במחשב המארח של השרת, וכאשר יידרש, השרת ישלוח את מידע זה וייבצע עליו פעולות.

שימוש זה אמנם שומר בזיכרון המחשב המארח את הטבלה, אך מפני שכל תהליך המבצע את אלגוריתם Map-Reduce יצטרך להשתמש בטבלה זו באופן בלתי תלוי בשרת עצמו, העתקה הייתה מבצעת בכל מקרה. לכן, כתוצאה מכך, בחרנו לשמור את הנתונים כ-hard copy, על מנת לא להכביד על תהליך שרת ה-gen_server, בשמירת הטבלה ב-server_state וכן לחסוך זיכרון מזיכרון ה-RAM.

בכל טבלה יופיעו **רשומות של סרטים** המכילים את המידע הבא (מוגדר כ-record בארלנג):

```
movie_data = {id, title, original_title, year, date_published, genre, duration, country,
language, director, writer, production_company, actors, description, avg_vote, votes,
budget, usa_gross_income, worldwide_gross_income, metacore, reviews_from_users,
evIEWS_from_critics}
```

מאחר וכל רשומה מכילה הרבה פרמטרים, השימוש הנ"ל עוזר לנו בניתוח המידע באופן ישיר. את המידע הזה אנו שולחים מהשרתים למאסטר ומהמאסטר ללקוח.

את **השאלות** אנו שולחים במערכת ב-record אחר המכיל את המידע הבא:

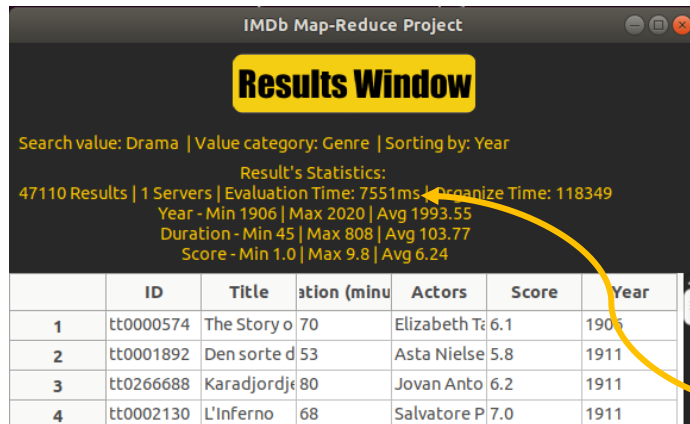
```
query = {type, searchVal, searchCategory, resultCategory}
```

כאשר ה-type נועד לשם להרחיב את אופציות המענה לשאלות, searchVal הוא הערך המוזן מהמשתמש כ-string, searchCategory הוא הפרמטר אותו אנו מחפשים (אותו חיפשנו ע"י ה-string) ולבסוף resultCategory הוא למעשה מסוג movie_data המציג לנו במערך אינדקציה (true/false) אילו קטגוריות לשלוח בחזרה.

ביצועים

ביצועי המערכת נמדדים בזמן הערכת המידע. יש לשים לב שמכיוון שאנו משתמשים בממשק המציג את כל המידע, זמן ייקח עד שהוא יוצג לפנינו כטקסט.

- בדקנו את המערכת כאשר לרשותך 5 שרתים מול שרת בודד, התוצאות מוצגות כאן:



IMDb Map-Reduce Project

Results Window

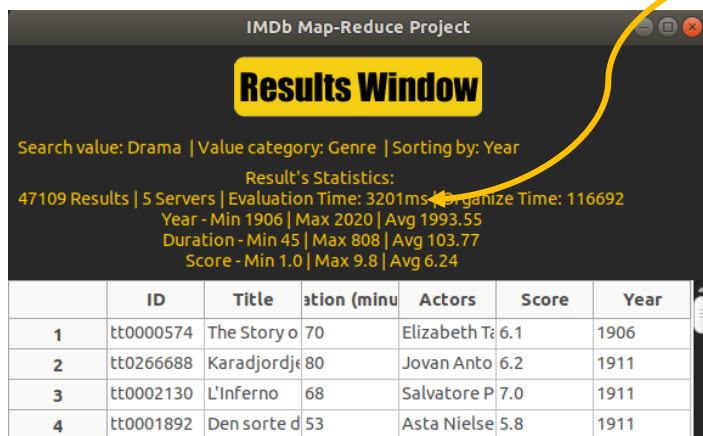
Search value: Drama | Value category: Genre | Sorting by: Year

Result's Statistics:
 47110 Results | 1 Servers | Evaluation Time: 7551ms | Organize Time: 118349
 Year - Min 1906 | Max 2020 | Avg 1993.55
 Duration - Min 45 | Max 808 | Avg 103.77
 Score - Min 1.0 | Max 9.8 | Avg 6.24

	ID	Title	ation (minu	Actors	Score	Year
1	tt0000574	The Story o	70	Elizabeth T	6.1	1906
2	tt0001892	Den sorte d	53	Asta Nielse	5.8	1911
3	tt0266688	Karadjordje	80	Jovan Anto	6.2	1911
4	tt0002130	L'Inferno	68	Salvatore P	7.0	1911

ניתן לראות כי עבור ביצוע של שרת בודד קיבלנו תוצאות גרועות. עבור 47k תוצאות לקח למערכת 7.5 שניות ומעלה לעבד את המידע.

זמן השגת המידע



IMDb Map-Reduce Project

Results Window

Search value: Drama | Value category: Genre | Sorting by: Year

Result's Statistics:
 47109 Results | 5 Servers | Evaluation Time: 3201ms | Organize Time: 116692
 Year - Min 1906 | Max 2020 | Avg 1993.55
 Duration - Min 45 | Max 808 | Avg 103.77
 Score - Min 1.0 | Max 9.8 | Avg 6.24

	ID	Title	ation (minu	Actors	Score	Year
1	tt0000574	The Story o	70	Elizabeth T	6.1	1906
2	tt0266688	Karadjordje	80	Jovan Anto	6.2	1911
3	tt0002130	L'Inferno	68	Salvatore P	7.0	1911
4	tt0001892	Den sorte d	53	Asta Nielse	5.8	1911

לעומת זאת, עבור 5 שרתים, ניתן לראות כי הביצועים השתפרו לפחות ב-4 שניות. למרות שהשיפור אינו לינארי בכמות השרתים, ניתן לראות כי הדבר אפקטיבי.

- כאשר שרת מתנתק או מתחבר, המאסטר מזהה זאת ומבצע ביזור מידי מחדש:

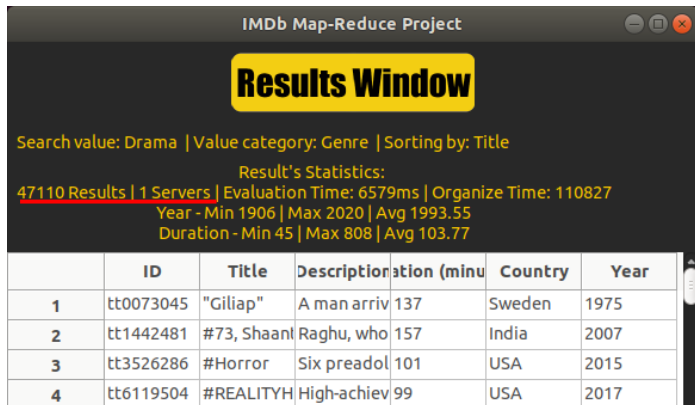
```
Eshell V11.0.2 (abort with ^G)
(master@master)1> master:start_link().
Master started.
send to server5@master, lines 2 to 17172.
send to server4@master, lines 17173 to 34343.
send to server3@master, lines 34344 to 51514.
send to server2@master, lines 51515 to 68685.
send to server1@master, lines 68686 to 85856.
Sent 85855 movie records to 5 servers in 87915 ms.
Message received: {nodeup,server1@master}
Message received: {nodeup,server2@master}
Message received: {nodeup,server3@master}
Message received: {nodeup,server4@master}
Message received: {nodeup,server5@master}
true
A node is down: server4@master
send to server1@master, lines 2 to 85856.
Sent 85855 movie records to 1 servers in 26981 ms.
```

בדוגמה הבאה, חיברנו 5 שרתים אל המאסטר ולאחר מכן ניתקנו 4 מהם וקיבלנו שינוי במאסטר.

רישום מידע ל-5 השרתים

שליחת המידע לשרת הבודד שנשאר

- כאשר שרת מתנתק בזמן בקשת תשובה, המאסטר מחזיר ללקוח תשובה רק מהשרתים שהוא הספיק להשיג מהם את המידע. לאחר מכן, המאסטר יהיה לא זמין, לשם ביזור מחדש – כך לא נאבד מידע ונוכל כבר בתשאול הבא לקבל את כלל התוצאות.



IMDb Map-Reduce Project

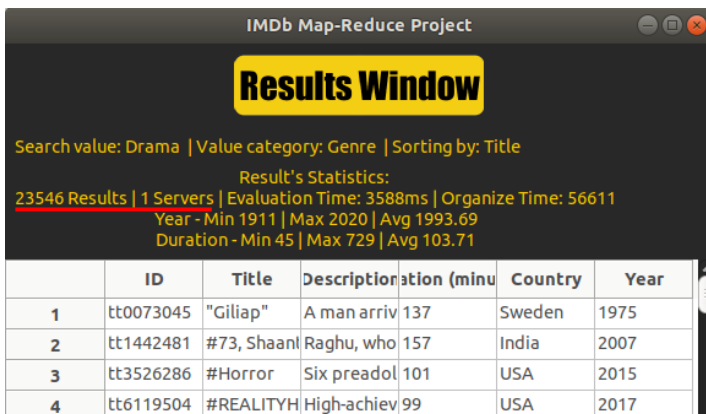
Results Window

Search value: Drama | Value category: Genre | Sorting by: Title

Result's Statistics:
47110 Results | 1 Servers | Evaluation Time: 6579ms | Organize Time: 110827
 Year - Min 1906 | Max 2020 | Avg 1993.55
 Duration - Min 45 | Max 808 | Avg 103.77

	ID	Title	Description	ation (minu	Country	Year
1	tt0073045	"Giliap"	A man arriv	137	Sweden	1975
2	tt1442481	#73, Shaanl	Raghu, who	157	India	2007
3	tt3526286	#Horror	Six preadol	101	USA	2015
4	tt6119504	#REALITYH	High-achiev	99	USA	2017

כאן ניתן לראות שבעבור שרת אחד שמכיל את כל המידע, חזרו כ-47k תוצאות.



IMDb Map-Reduce Project

Results Window

Search value: Drama | Value category: Genre | Sorting by: Title

Result's Statistics:
23546 Results | 1 Servers | Evaluation Time: 3588ms | Organize Time: 56611
 Year - Min 1911 | Max 2020 | Avg 1993.69
 Duration - Min 45 | Max 729 | Avg 103.71

	ID	Title	Description	ation (minu	Country	Year
1	tt0073045	"Giliap"	A man arriv	137	Sweden	1975
2	tt1442481	#73, Shaanl	Raghu, who	157	India	2007
3	tt3526286	#Horror	Six preadol	101	USA	2015
4	tt6119504	#REALITYH	High-achiev	99	USA	2017

לעומת זאת, עבור המקרה המתואר, עבדנו עם 2 שרתים בזמן ביצוע השאילתה. כאשר שלחנו אותה, הפלנו את אחד השרתים והמתנו לתשובה. סך התשובות הן כ-23k וזאת בגלל שאלו התשובות שהגיעו רק משרת בודד.

מסקנות

- כאשר מפתחים מערכת מבוזרת ומקבילית חשוב לתכנן מראש את הארכיטקטורה בצורה מפורטת שתמפה עבורנו את כלל הבעיות הנפרסות בפנינו.
- חשוב לשים לב למקרים בהם ייתכנו צווארי בקבוק במערכת – המאסטר למשל מנהל את כל התקשורת בין השרתים ובין הלקוחות, לכן חשוב לשים לב לביצוע הטיפול בבקשות בצורה יעילה ובלתי תלויה אחד בשני.
- גודל המידע ומספר השרתים לביזור משפיעים על תוצאות ביצוע השאילתה. ככל שנגדיל את מספר השרתים כך זמן התגובה יהיה קטן יותר. (*** לצרף תמונה של הניסוי ***)
הערה: בגלל מגבלות טכניות לא יכלנו לבדוק את המערכת עם כמה סרברים הממוקמים במחשבים שונים, אך ע"פ התיאוריה הדבר היה נותן ביצועים מצויינים.
- ארלנג כשפת תכנות נותנת כלים לתכנות מקבילי ומבוזר בצורה קלה ויעילה, לכן על מנת ליצור מערכת מבוזרת המעבדת כמויות גדולות של מידע מכמה יעדים מרוחקים, כדאי להשתמש בה כשפת התכנות הראשית.

Instructions

- **Compilation**

Need to compile all project files. Do it by writing the following command to the terminal:

```
> erl
> c(master). c(server). c(wxclient). c(parse_csv). c(dataDistributor).
```

- **System adjustments**

Insert the nodes that you will set on your system: servers in "*serverslist.txt*", master and clients in "*clientslist.txt*" (in that order!)

serverlists.txt	clientslists.txt	
server1@Host/IP	master@Host/IP	example: IP = 192.168.1.101
server2@Host/IP	client1@Host/IP	Host = ubuntu
server#@Host/IP	client#@Host/IP	

- **Run the system**

To run the system, you need at least 1 server and 1 client, and single master (always).

Server: `> erl -name server# -setcookie a -run server start_link`

Client: `> erl -name client# -setcookie a -run wxclient start`

Master: `> erl -name master -setcookie a -run master start_link`

*for client you may need to open shell "erl -name client -setcookie a", and write "wxclient:start()"

נספחים

- [סרטון הדגמה ב-YouTube](#)
- [הפרויקט ב-GitHub](#)