

2023 Spring Capstone

Midterm report

Team members : Huai-Che Chang, Layne Pu,
Wan-Yi Mao, Wen-Fang Lu

Agenda

- **Background**
- **Project Objectives**
- **Project scope and deliverables**
- **Front-end application**
 - Current UI
- **System overview**
- **Data flow**
- **Sentiment Analysis**
- **Challenge**

Team Member

Team Member	Role
Huai-Che Chang	Software Developer, Product Manager
Layne Pu	Team Leader, Software Developer
Wan-Yi Mao	Software Developer, DevOps
Wen-Fang Lu	Software developer, Scrum Master

Background

Financial Social N Informa

CLAUDIA BIANCOTTI, PA

reference [Carnegie](#)



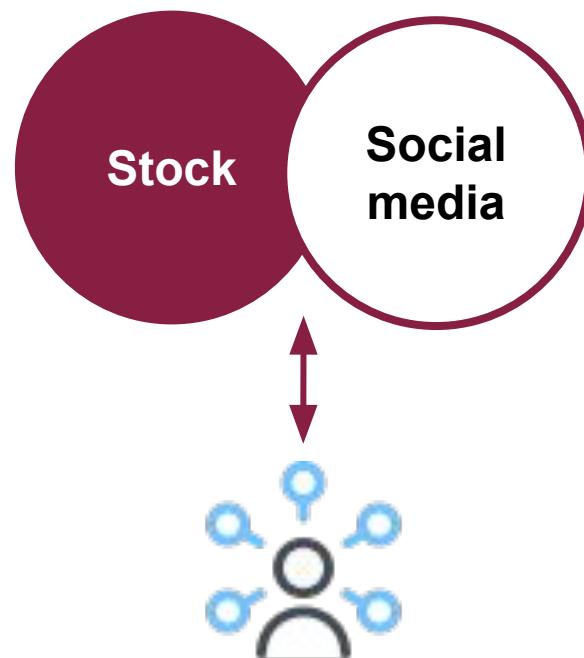
Unsplash © Patrick Weissenberger

From GameStop To Tesla – How social media is driving the stock market

reference [Maddyness news](#)

Project Objective

- Discover and analyze social media impact on stock market trend
- Provide an overview of public view about a specific stock



Project Scope And Deliverable

- Design - Project definition, System design document
- Development - Web application, Real-time data pipeline, Search engine, Sentiment analysis
- Testing - Unit testing, Functional testing, Integration testing, Load testing, Recovery testing
- Deployment - VT Cloud

Project Schedule - On Time !!

	Week 1-4	Week 5-6	Week 7-8	Week 9-10	Week 11-12	Week 13-14	Week 15-16
Project ideation	Red						
Schema definition	Orange	Orange					
Data streaming		Yellow	Yellow				
Web development			Green	Green	Green		
Deployment						Blue	Blue

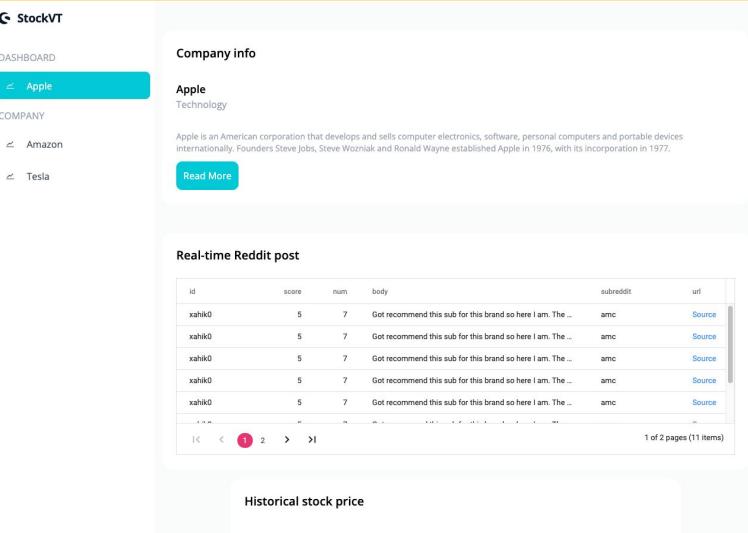
Front-end application

- A responsive and interactive web application using React, Syncfusion, Tailwindcss
- Expected features:
 - Company information
 - Historical data from yFinance API
 - Real-time posts from Reddit API (Highlight)

 **React**

 tailwindcss

 Syncfusion



The screenshot displays a responsive front-end application interface. On the left, a sidebar menu includes 'DASHBOARD' (highlighted in teal), 'COMPANY' (with 'Apple' selected), 'Amazon', and 'Tesla'. The main content area shows 'Company info' for Apple (Technology) with a brief description and a 'Read More' button. Below this is a section titled 'Real-time Reddit post' featuring a table of 5 rows with columns for id, score, num, body, subreddit, and url. The table shows identical entries for 'xahik0'. At the bottom, a 'Historical stock price' section is visible.

id	score	num	body	subreddit	url
xahik0	5	7	Got recommend this sub for this brand so here I am. The ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The ...	amc	Source

Responsive Front-end web application

8

Current UI



DASHBOARD

⌘ Apple

COMPANY

⌘ Amazon

⌘ Tesla

Company information



Company info

Apple

Technology

Apple is an American corporation that develops and sells computer electronics, software, personal computers and portable devices internationally. Founders Steve Jobs, Steve Wozniak and Ronald Wayne established Apple in 1976, with its incorporation in 1977.

Read More

Real-time Reddit post

id	sc...	num	body	subreddit	url
xahik0	5	7	Got recommend this sub for this brand so here I ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I ...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I ...	amc	Source
...

« < 1 2 > »

1 of 2 pages (11 items)

Real-time posts

Current UI

StockVT

DASHBOARD

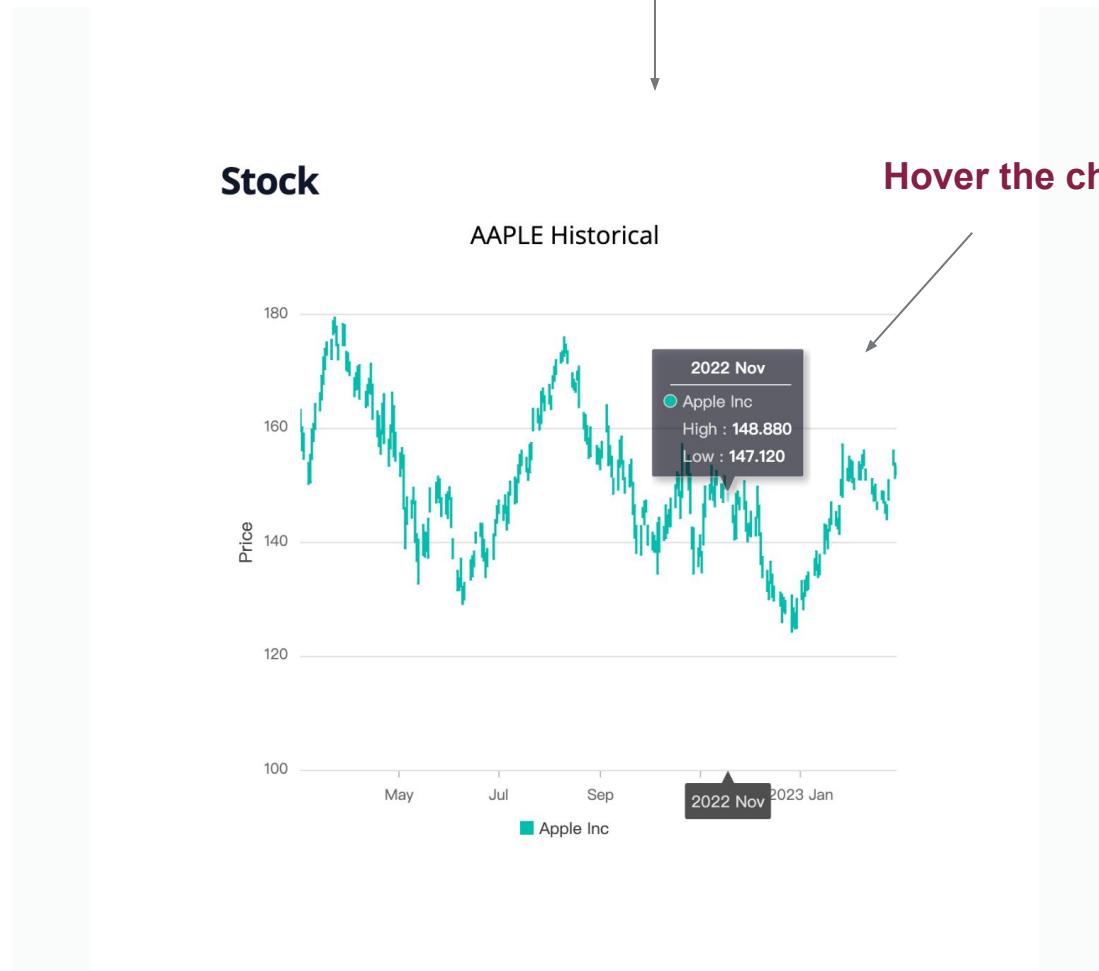
Apple

COMPANY

Amazon

Tesla

Stock price from the past 12 months



Current UI

Support different screen sizes for devices : ipad air, iphone 12 Pro, Web Browser (from left to right)

Company info

Apple
Technology

Apple is an American corporation that develops and sells computer electronics, software, personal computers and portable devices internationally. Founders Steve Jobs, Steve Wozniak and Ronald Wayne established Apple in 1976, with its incorporation in 1977.

[Read More](#)

Real-time Reddit post

id	s...	n...	body	subreddit	url
xahik0	5	7	Got recommend this sub for this brand so he...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so he...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so he...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so he...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so he...	amc	Source

1 < 1 2 > 11 1 of 2 pages (11 items)

Historical stock price

Stock

AAPLE Historical

Apple
Technology

Apple is an American corporation that develops and sells computer electronics, software, personal computers and portable devices internationally. Founders Steve Jobs, Steve Wozniak and Ronald Wayne established Apple in 1976, with its incorporation in 1977.

[Read More](#)

Real-time Reddit post

id	s...	n...	body
xahik0	5	7	Got recommend
xahik0	5	7	Got recommend
xahik0	5	7	Got recommend
xahik0	5	7	Got recommend

1 < 1 2 > 11 1 of 2 pages

Historical stock price

apple
Technology

Apple is an American corporation that develops and sells computer electronics, software, personal computers and portable devices internationally. Founders Steve Jobs, Steve Wozniak and Ronald Wayne established Apple in 1976, with its incorporation in 1977.

[Read More](#)

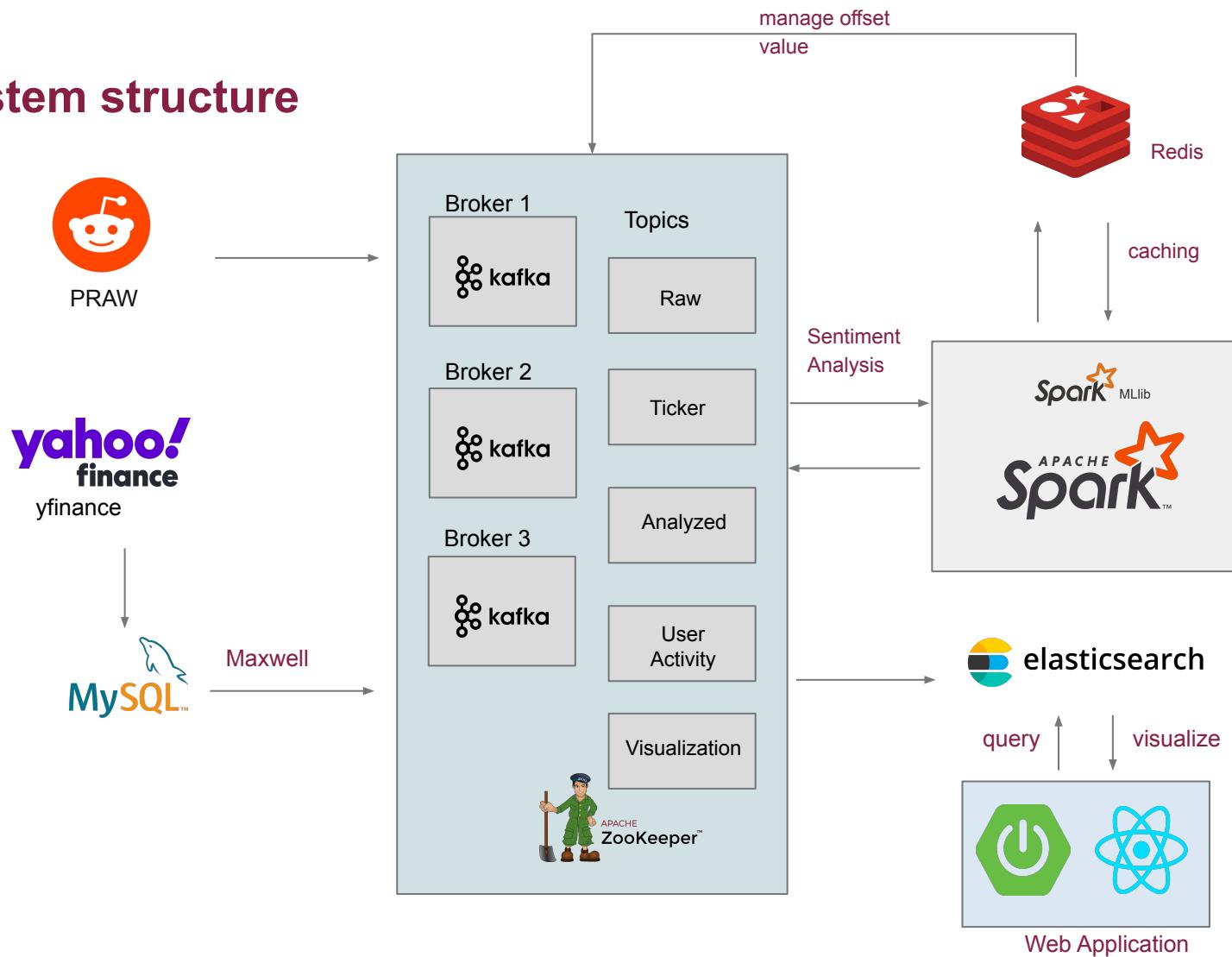
Real-time Reddit post

id	score	num	body	subreddit	url
xahik0	5	7	Got recommend this sub for this brand so here I am. The car is a '64 a...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The car is a '64 a...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The car is a '64 a...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The car is a '64 a...	amc	Source
xahik0	5	7	Got recommend this sub for this brand so here I am. The car is a '64 a...	amc	Source

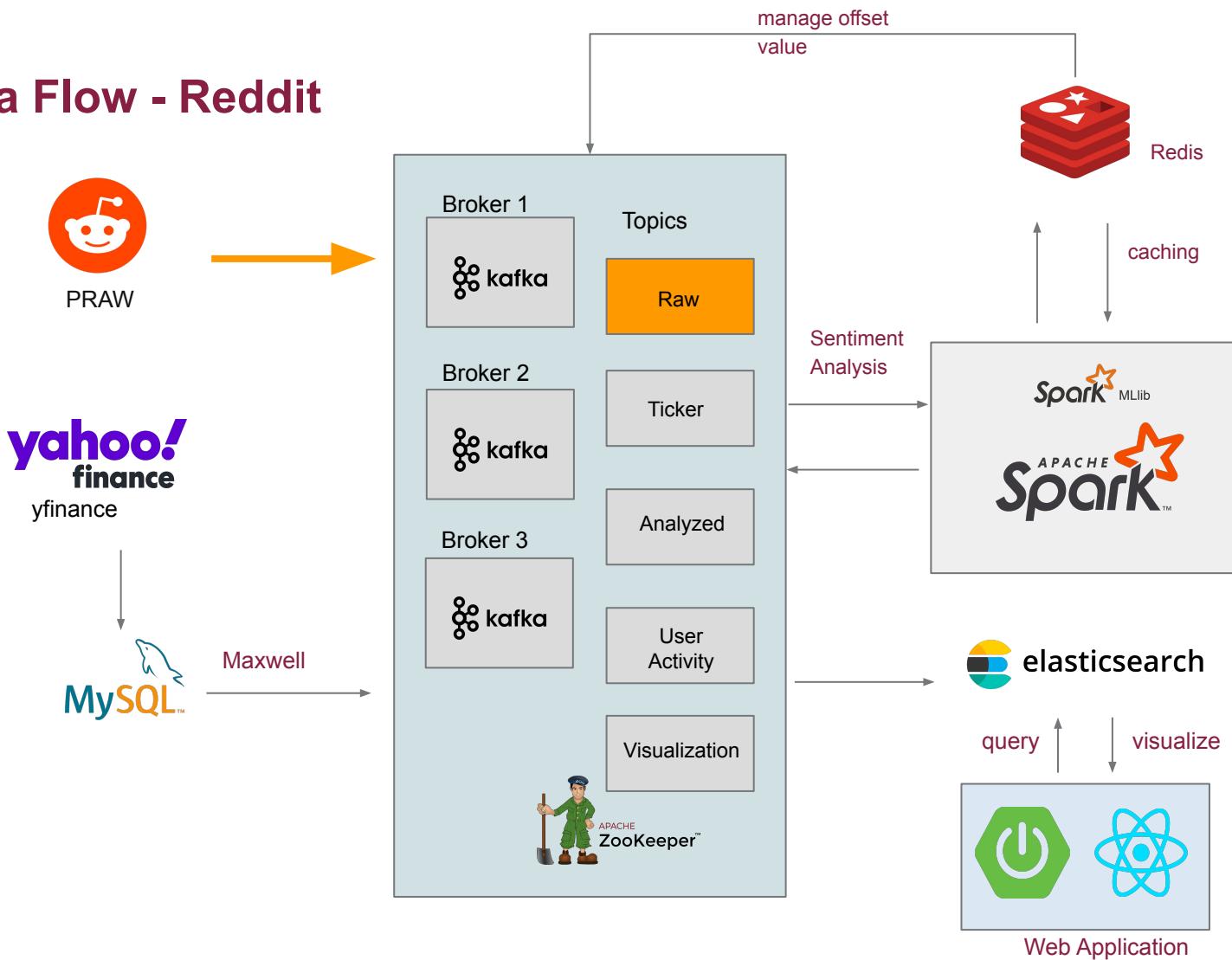
1 < 1 2 > 11 1 of 2 pages (11 items)

Historical stock price

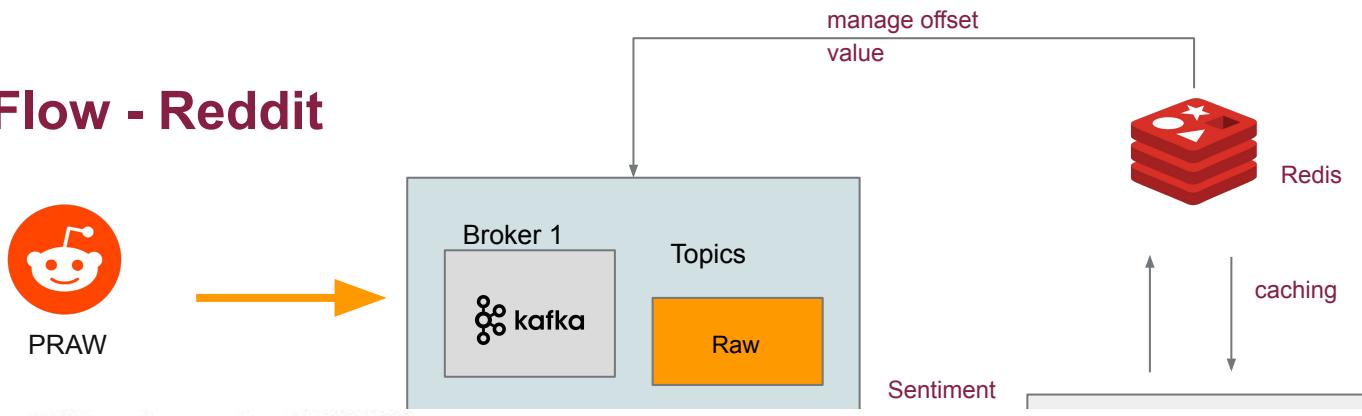
System structure



Data Flow - Reddit

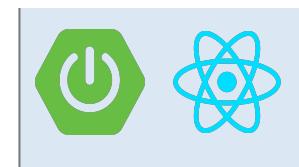


Data Flow - Reddit



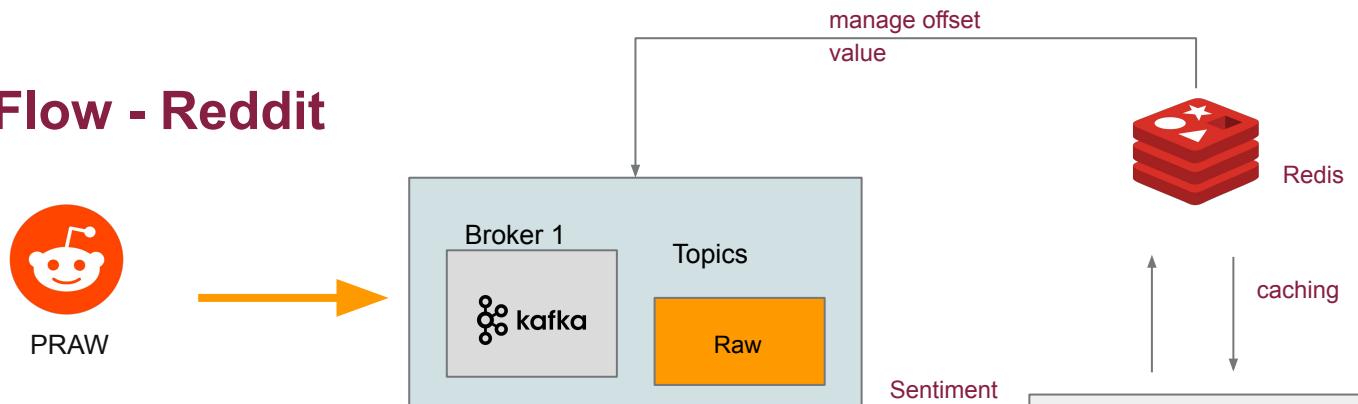
Subreddits - targeted (WIP)

name	ticker symbol	subreddits
AMC	AMC	"amc", "amcstock", "AMCSTOCKS"
GameStop	GME	"GME", "GMEJungle", "gme_meltdown", "gmeamcstonks", "gme_capitalists"
Tesla	TSLA	"Tesla", "teslainvestorsclub", "RealTesla", "teslamotors", "SpaceX", "elonmusk", "electricvehicles", "investing", "stocks", "wallstreetbets"



Web Application

Data Flow - Reddit



Subreddits - targeted (WIP)

name	ticker symbol	subreddits
A	125 # Start threads to handle subreddit updates 126 amc_thread = threading.Thread(target=handle_subreddit_updates, args=(amc_subreddits, AMC_TOPIC)) 127 gme_thread = threading.Thread(target=handle_subreddit_updates, args=(gme_subreddits, GME_TOPIC)) 128 tsla_thread = threading.Thread(target=handle_subreddit_updates, args=(tsla_subreddits, TSLA_TOPIC)) 129 amc_thread.start() 130 gme_thread.start() 131 tsla_thread.start()	
G	132 133 # Wait for threads to complete 134 amc_thread.join() 135 gme_thread.join() 136 tsla_thread.join()	
T		

Reddits posts

- Hot/ Top/ Controversial/ New
- Rising

↑ Posted by u/kickass_bacon 8 years ago ☀

7 What is the difference between new,rising, hot, top, and controversial posts ?

↓ another related question: how to know the trending/active posts so i can contribute to ?

2 Comments Award Share Save ...

This thread is archived
New comments cannot be posted and votes cannot be cast

Sort By: Best ▾

 Redditnotes · 8 yr. ago · edited 8 yr. ago

Hot is what's been getting a lot upvotes/comments recently

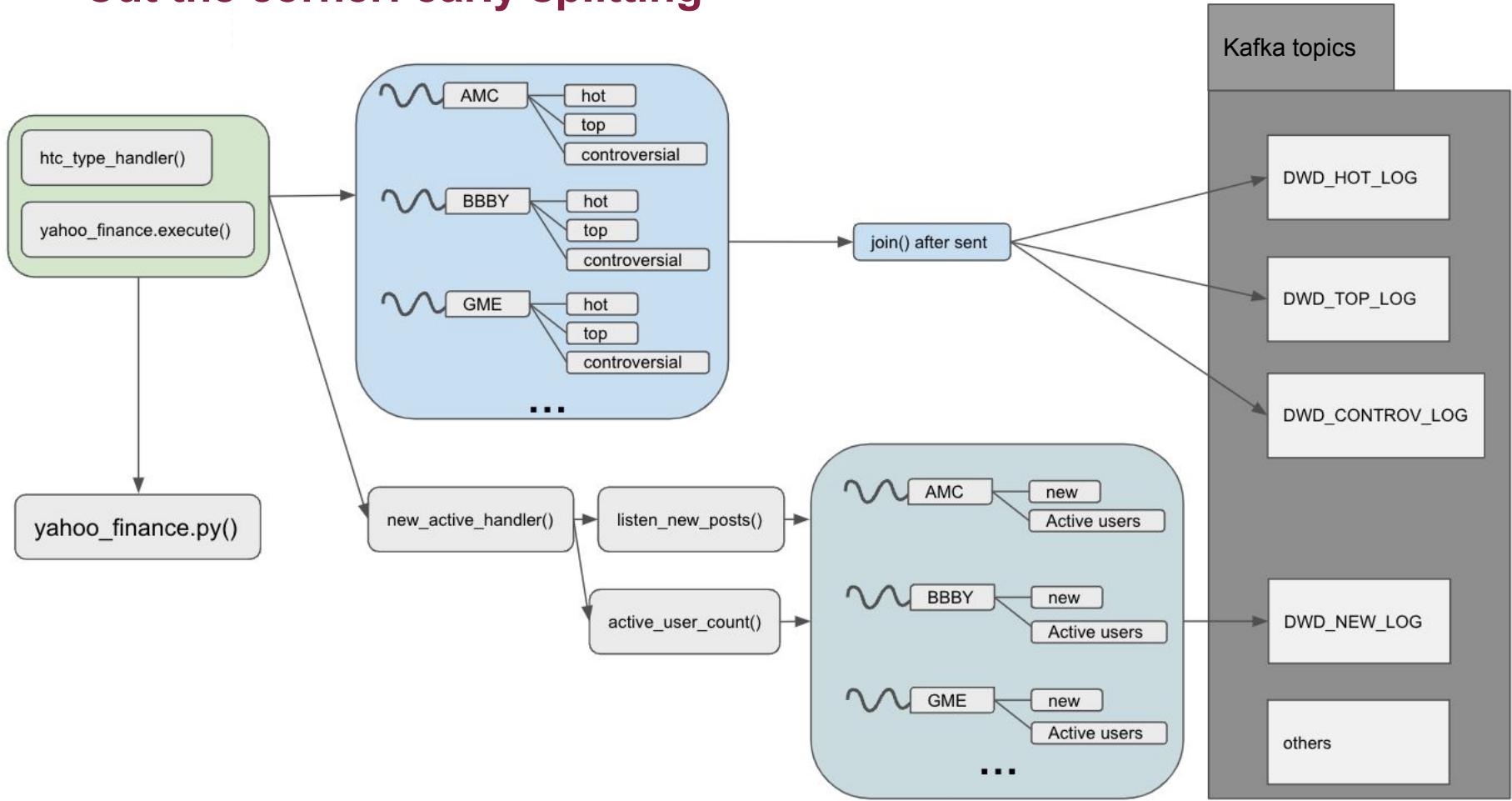
New sorts post by the time of submission with the newest at the top of the page. It sorts new posts in the area of reddit you clicked the 'new' tab. So all.reddit.com/new gives you the latest post from the entirety (almost) of reddit. While clicking new in askreddit will give you their latest posts. This applies to all the tabs I explain below.

Rising is what is getting a lot of activity (comments/upvotes) right now. This is the category you are looking for.

Controversial is what's getting multiple downvotes and upvotes.

Top is what has gotten the most upvotes over the set period.

Cut the corner: early splitting



Topic Messages: DWD_HOT_LOG

First Offset: 0 Last Offset: 219 Size: 219

Topic Messages: DWD_TOP_LOG

First Offset: 0 Last Offset: 230 Size: 230

Topic Messages: DWD_CONTROVERSIAL_LOG

First Offset: 0 Last Offset: 210 Size: 210

Topic Messages: DWD_NEW_LOG

First Offset: 0 Last Offset: 90 Size: 90

Partition 0

Offset 0

messages 100

Key format

DEFAULT

Message format

DEFAULT

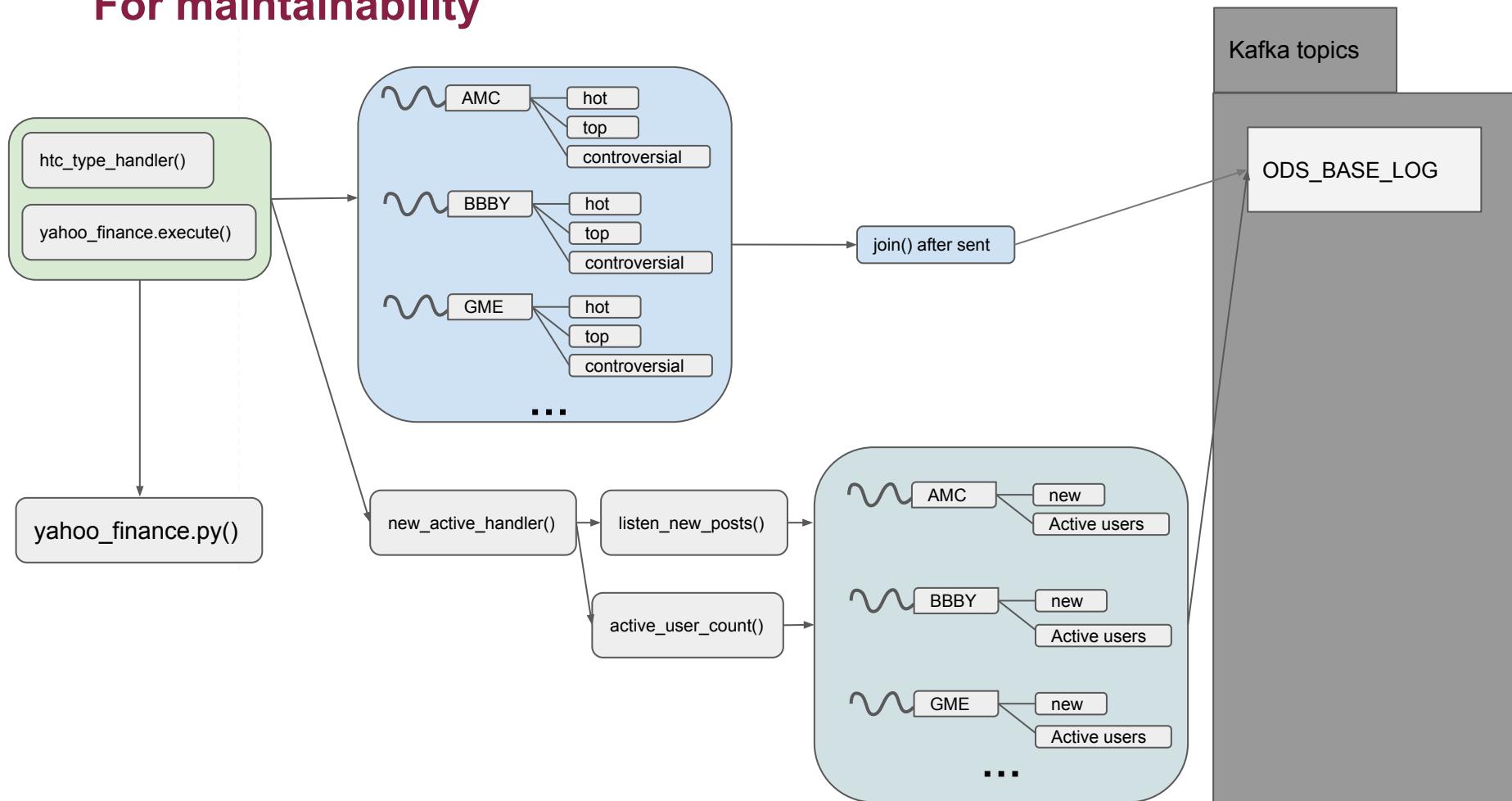
View Messages

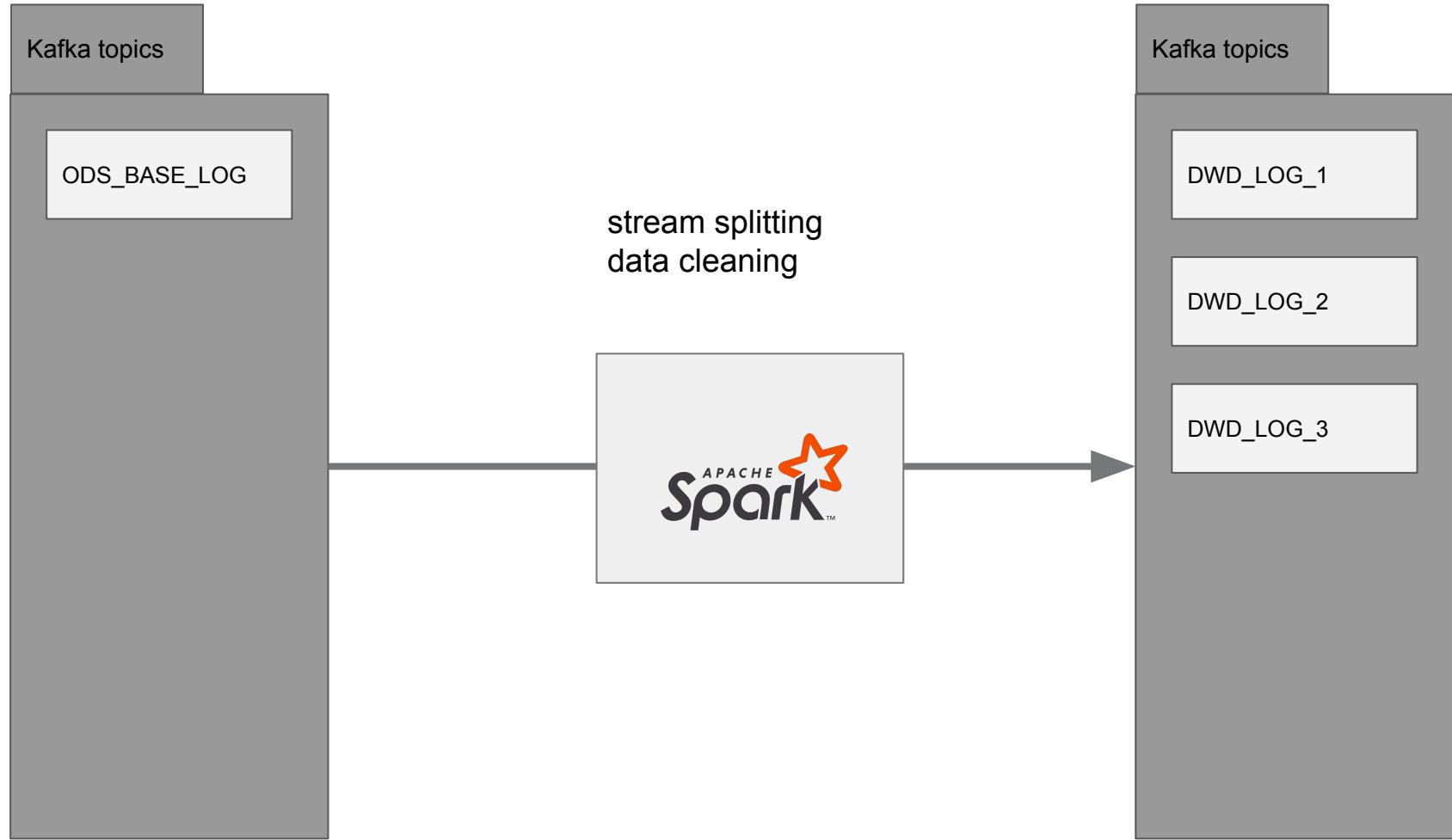
Offset: 0 Key: empty Timestamp: 2023-03-12 22:57:44.006 Headers: empty

{

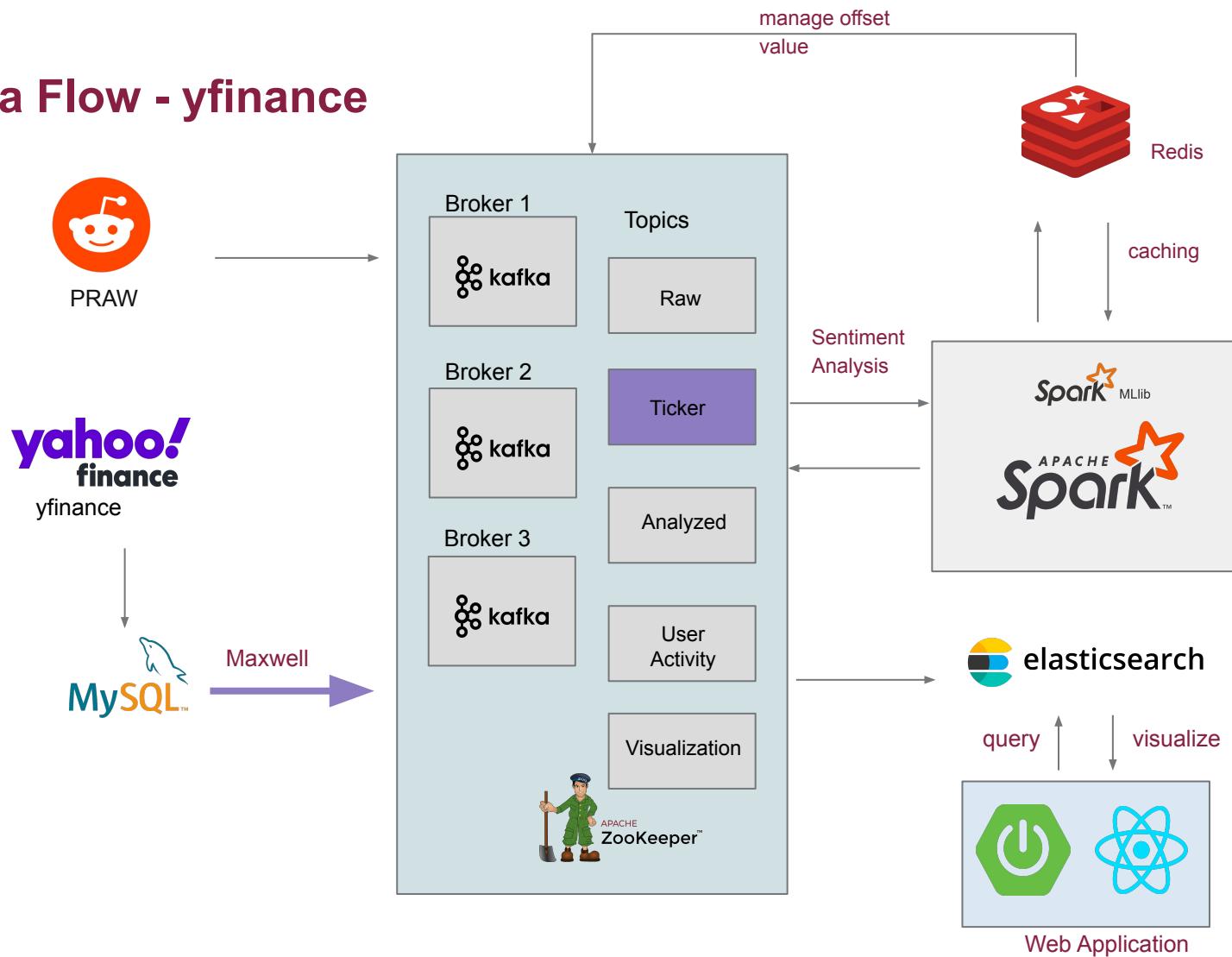
```
"author": "newwobblywheeler",
"author_flair_text": null,
"created_time": "2023-03-08 11:06:08.000000Z",
"id": "11m06zr",
"is_original": false,
"is_self": true,
"permalink": "/r/BB_Stock/comments/11m06zr/blackberry_drop_yesterday/",
"title": "Blackberry Drop yesterday",
"body": " It seem that yesterday the short volume was \~6 million shares and today \~ 7.7million shares are available for shorting",
"score": 0,
"upvote_ratio": 0.38,
"num_comments": 1,
"url": "https://www.reddit.com/r/BB_Stock/comments/11m06zr/blackberry_drop_yesterday/",
"subreddit": "BB_stock",
"ticker": "BB",
"label": "new"
}
```

For maintainability





Data Flow - yfinance



Data Retrieved from Yahoo - Historical stock market data

name	attribute	description
"trade_date"	history.Date	The date of a particular trading day.
"open"	history.Open	The opening price of the stock for the trading day.
"high"	history.High	The highest price the stock reached during the trading day.
"low"	history.Low	The lowest price the stock reached during the trading day.
"close"	history.Close	The closing price of the stock for the trading day.
"volume"	history.Volume	The total number of shares traded during the trading day.

Data Retrieved from Yahoo - Current stock market data

"market_cap"	fast_info.marketCap	The market capitalization of the financial instrument.
"day_open"	fast_info.open	The opening price of the financial instrument for the day.
"pre_close"	fast_info.previousClose	The previous day's closing price of the financial instrument.
"quote_type"	fast_info.quoteType	The type of financial instrument.
"reg_pre_close"	fast_info.regularMarketPreviousClose	The previous regular trading hours closing price of the financial instrument.
"num_shares"	fast_info.shares	The number of shares outstanding for the financial instrument.
"ten_day_aver_vol"	fast_info.tenDayAverageVolume	The ten-day average volume for the financial instrument.
"three_month_aver_vol"	fast_info.threeMonthAverageVolume	The three-month average volume for the financial instrument.

...and so on

Data Retrieved from Yahoo - Historical meta data

name	attribute	description
currency	history_metadata.currency	The currency in which the stock is priced.
symbol	history_metadata.symbol	The ticker symbol for the stock.
exchange_name	history_metadata.exchangeName	The name of the exchange on which the stock is traded.
instrument_type	history_metadata.instrumentType	The type of stock (e.g., common stock, preferred stock, ETF, etc.).
first_trade_date	history_metadata.firstTradeDate	The date and time of the first trade for the stock.
regular_market_time	history_metadata.regularMarketTime	The date and time of the most recent trade for the stock.
gmt_offset	history_metadata.gmtoffset	The GMT offset for the time zone in which the stock is traded.
timezone	history_metadata.timezone	The time zone in which the stock is traded.

...and so on

Data Retrieved from Yahoo - Others

- Historical share count data
- Major holders data
- Institutional and mutual fund holders' data
- Earnings data (EPS)

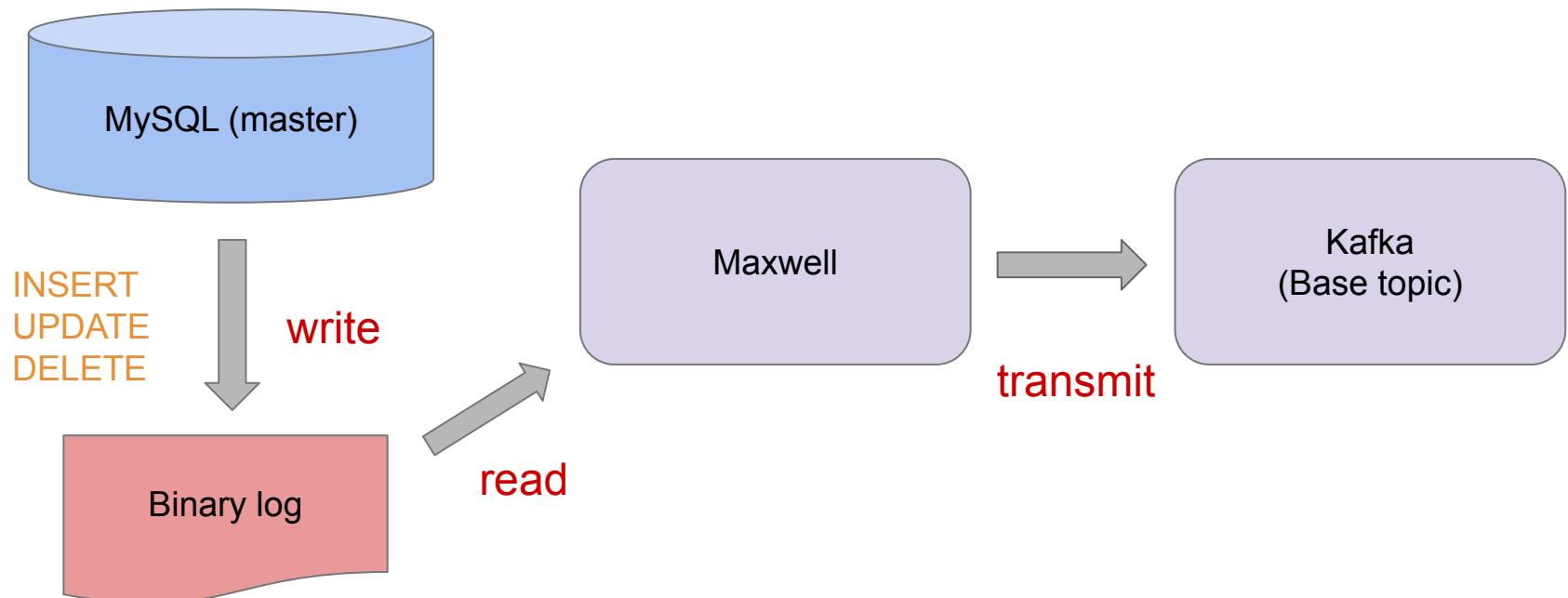
MySQL Database - Historical stock market data

#	symbol	date	open	high	low	close	volume
1136	AMC	2021-03-10	11.02	12.47	9.51	9.85	261918600
1137	AMC	2021-03-11	10.65	10.87	9.9	10.28	83933600
1138	AMC	2021-03-12	10.16	11.4	9.94	11.16	111146700
1139	AMC	2021-03-15	12.18	14.49	11.85	14.04	277713300
1140	AMC	2021-03-16	13.62	13.62	12.34	13.02	125967600
1141	AMC	2021-03-17	13.24	13.66	13	13.56	78053600
1142	AMC	2021-03-18	14.34	14.54	13.57	14	121418000
1143	AMC	2021-03-19	14.14	14.18	13.28	13.93	153206000
1144	AMC	2021-03-22	13.15	13.19	11.76	12.49	88760100
1145	AMC	2021-03-23	11.46	11.93	10.37	10.66	87923200
1146	AMC	2021-03-24	10.82	11.21	8.93	9.02	81850700
1147	AMC	2021-03-25	8.96	11.32	8.95	10.94	131192800
1148	AMC	2021-03-26	11.27	11.53	10.01	10.24	84633100
1149	AMC	2021-03-29	10.32	10.76	10.09	10.35	37330700
1150	AMC	2021-03-30	10.31	10.52	9.76	10.35	39020500
1151	AMC	2021-03-31	10.4	10.47	10.05	10.21	29832300
1152	AMC	2021-04-01	10.37	10.36	9.15	9.74	77677200

MySQL Database - Current stock market data

	id	symbol	currency	day_high	day_low	exchange	fifty_day_averag...	last_price	last_volume	market_cap	day_open	pre_close
▶	1	AMC	USD	5.92	5.56	NYQ	5.3902	5.65	16080500	2900314458	5.74	5.79
	2	GME	USD	18.38	17.48	NYQ	19.7676	17.52	2644900	5336206419	18.07	18.22
	3	TSLA	USD	185.18	172.51	NMS	165.4872	172.92	168837500	547136182806	180.25	181.375
	4	AAPL	USD	154.54	150.23	NMS	143.0146	150.59	53755100	2382619829326	153.56	153
	5	MSFT	USD	259.56	251.58	NMS	248.038	252.32	26620700	1888494125069	255.82	253.46
	6	GOOG	USD	95.92	92.355	NMS	94.5132	92.66	24409700	1184018766108	94.49	94.745
	7	AMZN	USD	96.21	92.18	NMS	95.105	92.25	56136400	945313433856	93.68	93.785
	8	BRK-A	USD	474834	460924	NYQ	469852.88	461030	4200	668097014200	471004	472002
	9	NVDA	USD	244.54	233.83	NMS	197.2022	234.36	50040900	584025121520	241.75	242.45
	10	TSM	USD	90.69	88.57	NYQ	88.2248	88.88	10347700	460938753402	89.71	90.73
	11	META	USD	188.93	180.3	NMS	156.8092	181.69	26540800	471056767929	186.35	184.85
	12	DIS	USD	100.05	95.92	NYQ	101.4828	96.14	8265300	175629510746	99.45	99.29
	13	ORCL	USD	89.64	86.53	NYQ	87.247	86.87	10506400	234223254634	88.33	88.5
	14	INTC	USD	26.98	25.97	NMS	27.8228	26.44	54863600	109382280517	25.98	25.92
	15	MA	USD	364.02	353.91	NYQ	365.2328	354.44	2414700	337867454082	359.25	359.03
	16	JPM	USD	137.35	129.22	NYQ	139.1984	130.34	25865300	385893710701	136.76	137.79
	17	WMT	USD	138.98	136.92	NYQ	142.7726	137.17	5137900	369920051062	138.83	138.2

Maxwell: MySQL Data Monitoring and Transmission to Kafka



- Incremental synchronization

Organize and Categorize data with different topics

Topic Messages: DWD_FAST_INFO_I

First Offset: 0 Last Offset: 75 Size: 75

Partition 0 Offset

Offset: 0 Key: empty Timestamp: 2023-03-10 11:06:15.342 Headers: empty

⌚ {"date": "2021-01-04", "volume": 40090000, "symbol": "GME", "high": 4.775, "low": 4.2875, "id": 1, "close": 4.3125, "open": 4.75}

Offset: 1 Key: empty Timestamp: 2023-03-10 11:06:15.342 Headers: empty

⌚ {"date": "2021-01-05", "volume": 19846000, "symbol": "GME", "high": 4.52, "low": 4.3075, "id": 2, "close": 4.3425, "open": 4.3375}

Offset: 2 Key: empty Timestamp: 2023-03-10 11:06:15.342 Headers: empty

⌚ {"date": "2021-01-06", "volume": 24224800, "symbol": "GME", "high": 4.745, "low": 4.3325, "id": 3, "close": 4.59, "open": 4.335}

Offset: 3 Key: empty Timestamp: 2023-03-10 11:06:15.342 Headers: empty

⌚ {"date": "2021-01-07", "volume": 24517200, "symbol": "GME", "high": 4.8625, "low": 4.505, "id": 4, "close": 4.52, "open": 4.6175}

Offset: 4 Key: empty Timestamp: 2023-03-10 11:06:15.342 Headers: empty

⌚ {"date": "2021-01-08", "volume": 25928000, "symbol": "GME", "high": 4.575, "low": 4.27, "id": 5, "close": 4.4225, "open": 4.545}

Offset: 5 Key: empty Timestamp: 2023-03-10 11:06:15.342 Headers: empty

⌚ {"date": "2021-01-11", "volume": 59632000, "symbol": "GME", "high": 5.1625, "low": 4.7525, "id": 6, "close": 4.985, "open": 4.8525}

Offset: 6 Key: empty Timestamp: 2023-03-10 11:06:15.343 Headers: empty

⌚ {"date": "2021-01-12", "volume": 28242800, "symbol": "GME", "high": 5.1, "low": 4.83, "id": 7, "close": 4.9875, "open": 4.99}

Offset: 7 Key: empty Timestamp: 2023-03-10 11:06:15.343 Headers: empty

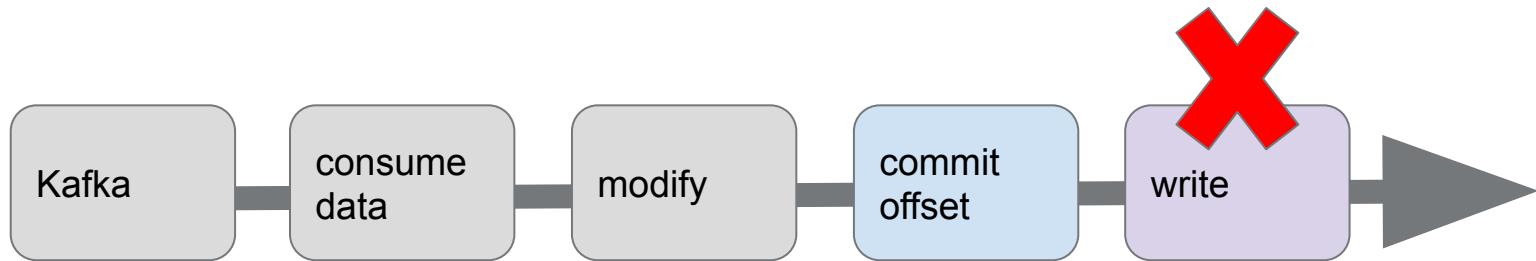
⌚ {"date": "2021-01-13", "volume": 578006800, "symbol": "GME", "high": 9.6625, "low": 5.0075, "id": 8, "close": 7.85, "open": 5.105}

Topic Messages: DWD_STOCK_HIST_I

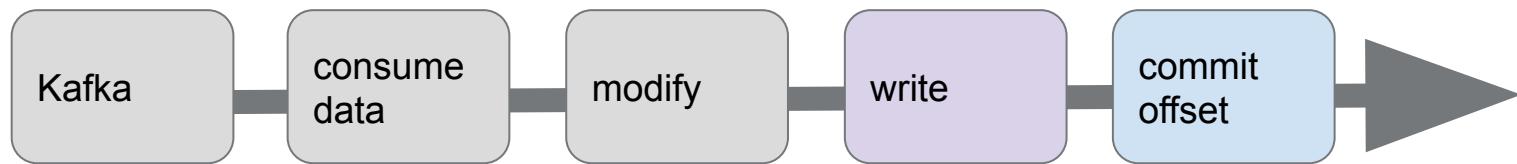
First Offset: 0 Last Offset: 1635 Size: 1635

Partition 0 Offset # messages 100 Key format DEFAULT Message format DEFAULT View

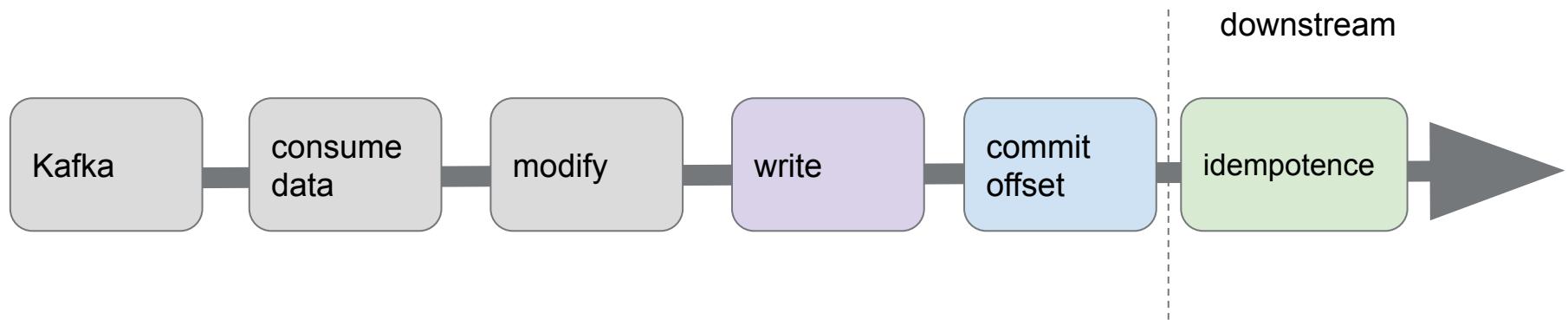
It's all for the consumer - data loss



It's all for the consumer - duplicated consumption



It's all for the consumer - solution



1. read offset from redis
2. use these offset values to consume kafka data
3. record the partition + end of offset
4. use the consumed data
5. submit the offset in step 3 to Redis

Save the offset to Redis

Question: Where is the offset from?

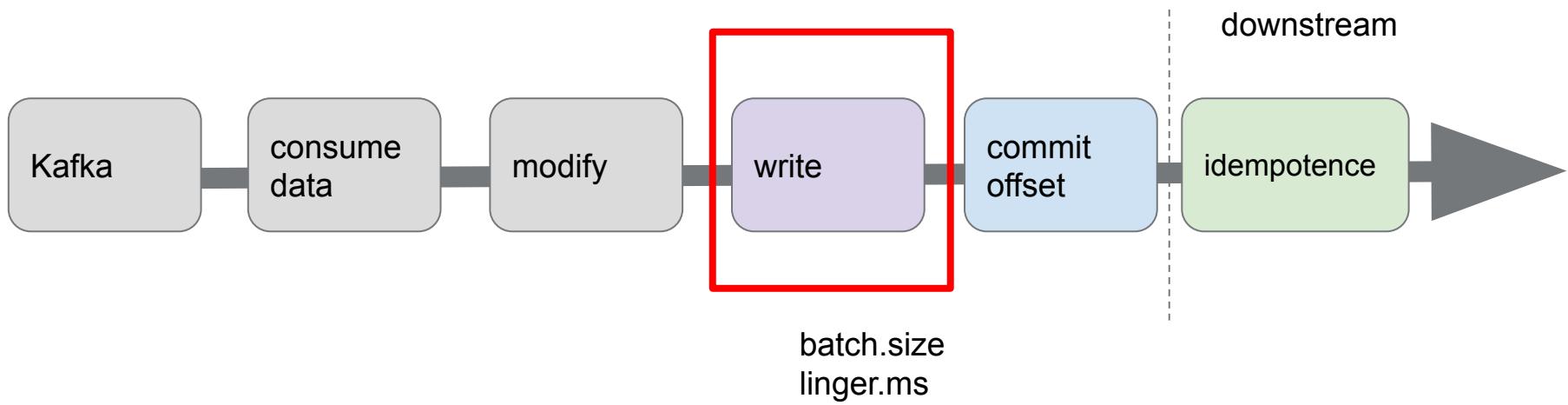
from the consumed data. The structure in Kafka is key -> val pair
key: group ID + topic + partition -> val: offset value

Question: How to save in Redis?

key: group ID + topic

value: partition + offset, partition + offset,

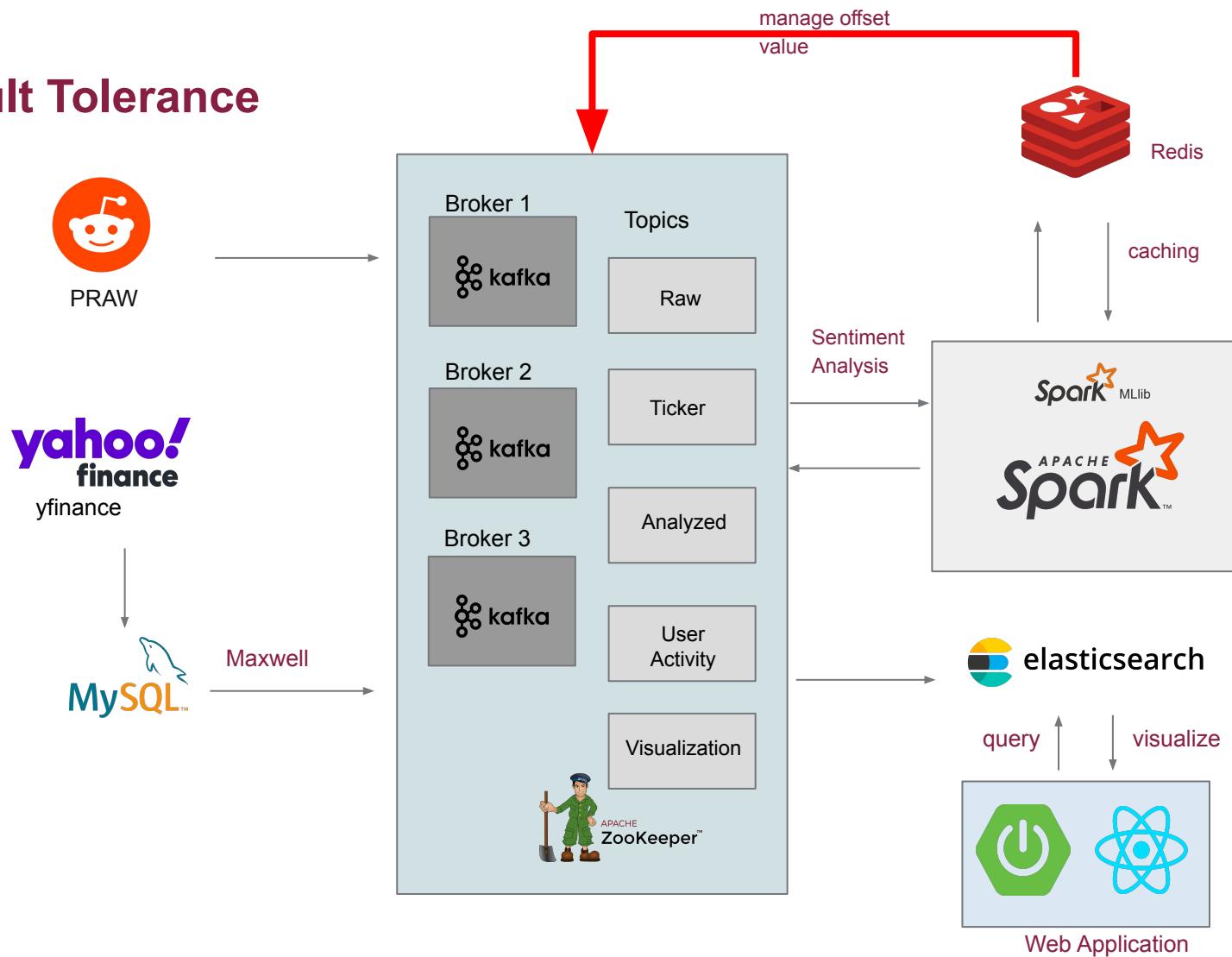
Inside the Kafka - still linger in cache?



solution1: send one msg after one synchronously (slow)

solution2: **flush** before commit offset. Flush for each partition

Fault Tolerance



Save the offset to Redis

1. run the spring boot application.

```
Successfully read offset: {0=2878, 1=4734, 2=2751, 3=3826}
Successfully submit offset: {0=2918, 1=4774, 2=2791, 3=3866}
Successfully submit offset: {0=2926, 1=4781, 2=2799, 3=3873}
Successfully submit offset: {0=2938, 1=4793, 2=2811, 3=3885}
Successfully submit offset: {0=2950, 1=4805, 2=2823, 3=3898}
Successfully submit offset: {0=2962, 1=4818, 2=2835, 3=3910}

Process finished with exit code 130 (interrupted by signal 2: SIGINT)
```

2. stop the application

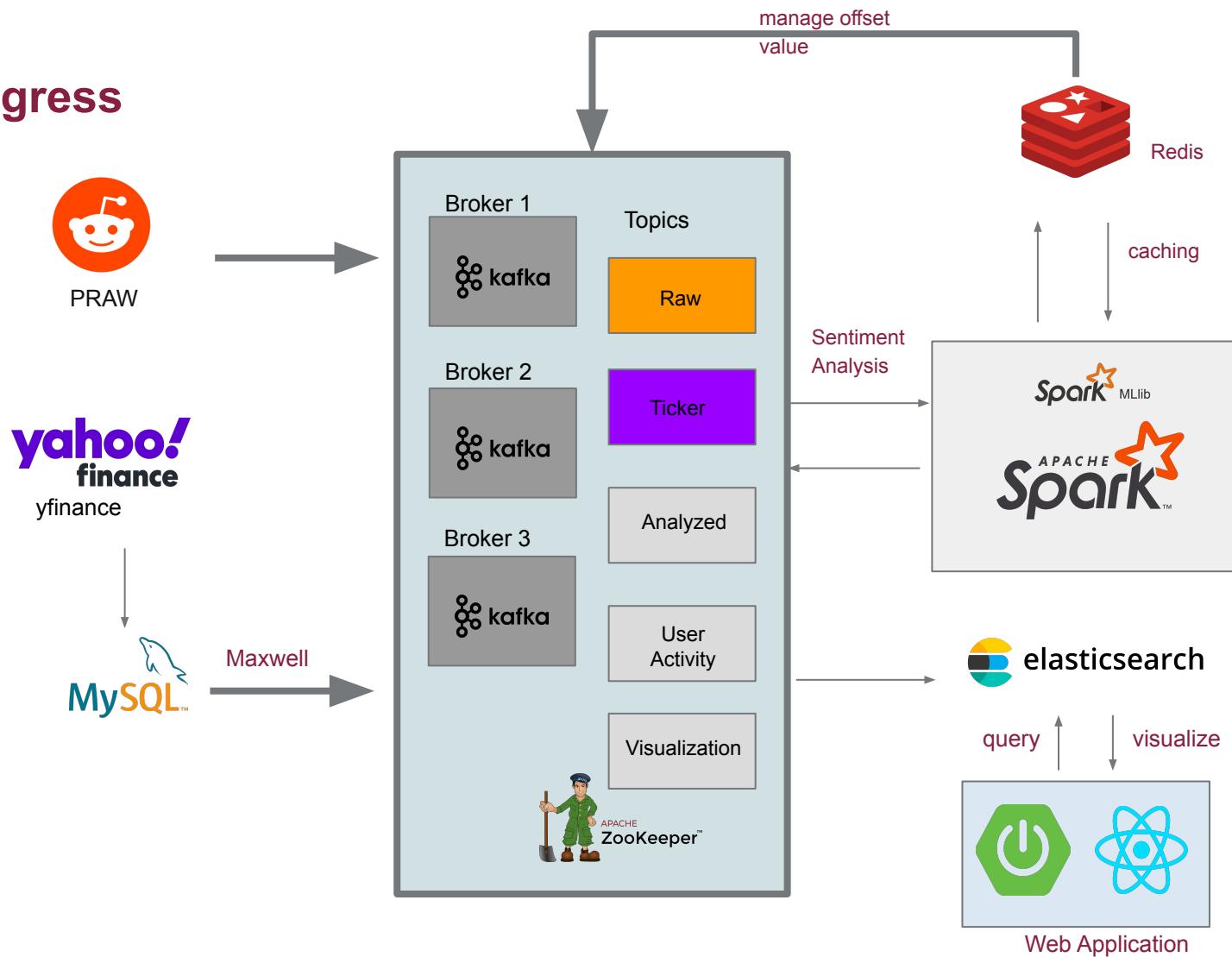
3. run the application again, check if read offset values is the same as when we stopped the application.

(0=2962, 1=4818, 2= 2835, 3=3910) 

```
/Users/roy/Library/Java/JavaVirtualMachines/corretto-1.8.0_362/Contents/Home/bin/java ...
Successfully read offset: {0=2962, 1=4818, 2=2835, 3=3910}
Successfully submit offset: {0=3012, 1=4867, 2=2885, 3=3960}
Successfully submit offset: {0=3023, 1=4878, 2=2896, 3=3970}
Successfully submit offset: {0=3035, 1=4890, 2=2908, 3=3982}

Process finished with exit code 130 (interrupted by signal 2: SIGINT)
```

Progress



Elasticsearch

- support the full-text search functionality within our front-end application

Console Search Profiler Grok Debugger Painless Lab BETA

History Settings Help

```
1 PUT /project/_doc/1
2 {
3   "id": 1,
4   "title": "Real-time monitoring of
      volatile, retail investor-drivenprice"
5   ,
6   "memberList": [
7     {"id":1,"name":"Huai-Che Chang"},
8     {"id":2,"name":"Wen-Fang Lu"}
9     {"id":3,"name":"Wan-Yi Mao"},
10    {"id":4,"name":"Layne Pu"}]
11
12
13
14
15 GET project/_search
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
```

Sentiment Analysis

- Collecting Reddit Posts and provide the sentiment analysis for specific stocks
- Real-time reddit posts analysis and history sentiment analysis result for data visualization
- Models:
 - TextBlob
 - Vader
 - BERT

	title	score	id	url	comms_num	created	body	timestamp	NLTK_compound	NLTK	Textblob
0	It's not about the money, it's about sending a...	55	I6ulcx	https://v.redd.it/6j75regs72e61	6	1.611863e+09	NaN	2021-01-28 21:37:41	Neutral	Neutral	Neutral
1	Math Professor Scott Steiner says the numbers ...	110	I6uibd	https://v.redd.it/ah50lyny62e61	23	1.611862e+09	NaN	2021-01-28 21:32:10	Negative	Negative	Neutral
2	Exit the system	0	I6uhhn	https://www.reddit.com/r/wallstreetbets/comment...	47	1.611862e+09	The CEO of NASDAQ pushed to halt trading "to g..."	2021-01-28 21:30:35	Neutral	Neutral	Neutral
3	NEW SEC FILING FOR GME! CAN SOMEONE LESS RETAR...	29	I6ugk6	https://sec.report/Document/0001193125-21-019848/	74	1.611862e+09	NaN	2021-01-28 21:28:57	Negative	Neutral	Negative
4	Not to distract from GME, just thought our AMC...	71	I6ufgy	https://i.redd.it/4h2sukb662e61.jpg	156	1.611862e+09	NaN	2021-01-28 21:26:56	Positive	Positive	Postive
5	WE BREAKING THROUGH	405	I6uf7d	https://i.redd.it/2wef8tc062e61.png	84	1.611862e+09	NaN	2021-01-28 21:26:30	Neutral	Neutral	Neutral

Sentiment Analysis - TextBlob, BERT

- Fine-tuning with large data of labeled stock-related texts
 - Financial Phrasebank
 - StockTwits

Sentiment Analysis - Vader

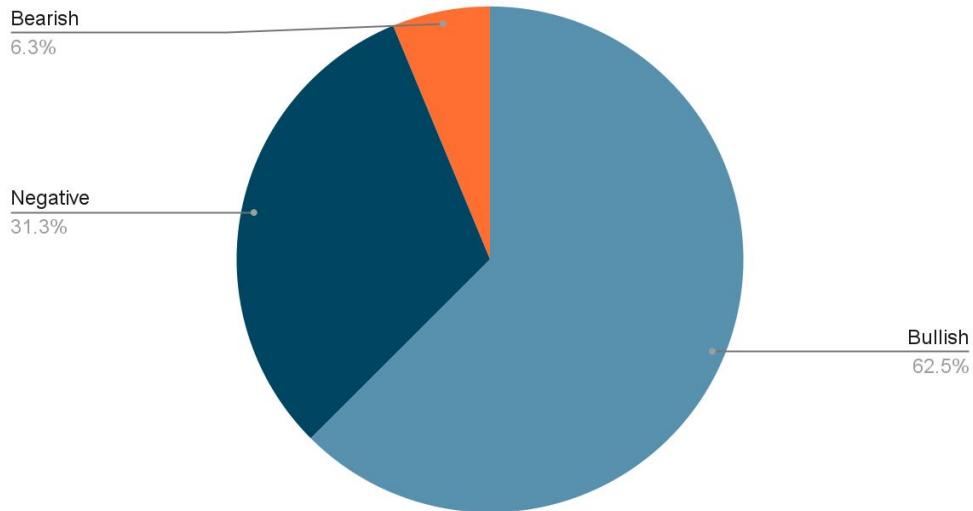
```
new_words = {  
    'citron': -4.0,  
    'hidenburg': -4.0,  
    'moon': 4.0,  
    'highs': 2.0,  
    'mooning': 4.0,  
    'long': 2.0,  
    'short': -2.0,  
    'call': 4.0,  
    'calls': 4.0,  
    'put': -4.0,  
    'puts': -4.0,  
    'break': 2.0,  
    'tendie': 2.0,  
    'tendies': 2.0,  
    'town': 2.0,  
    'overvalued': -3.0,  
    'undervalued': 3.0,  
}
```

stock	Bearish	Neutral	Bullish	Total_Compound
WISH	0.013	0.477	0.510	0.703
ON	0.041	0.594	0.366	0.574
CLNE	0.029	0.779	0.192	0.295
WEN	0.063	0.747	0.191	0.226
GOEV	0.029	0.788	0.184	0.177
CLOV	0.065	0.766	0.168	0.142
WKHS	0.019	0.828	0.153	0.221
GO	0.005	0.895	0.100	0.229
TLRY	0.037	0.866	0.097	0.098
AAPL	0.099	0.843	0.057	-0.209

[ref](#)

Sentiment Analysis - Result

Points scored



ID	Title	Body	TimeStamp	Result
xxae23	Currently Holding AMC and NOK - Is it retarded that I think I should move it all to GME today?	I believe right now is one of those rare opportunities that we all can help and do good. Some of these companies like GME, AMC are good companies that's been hit hard by this pandemic. Hedgefunds and Wallstreet just want..	2021-01-28 21:18:25	Neutral



Addressed problems

- Exactly-once semantic
- Kafka pageCache, flush
- Redis connection pool: Jedis
- Data synchronization (mysql)

Future Work

- Back-end
 - idempotency of rising posts
- Front-end
 - Connect to real-time posts from Reddit real-time API
 - Searching interface integrated with Elasticsearch
- Sentiment Analysis in Spark ML
- Sentiment Result store in Elasticsearch
- Implement Elasticsearch API to provide full-text search capabilities for our front-end application
- Gather additional valuable data from the yfinance API and store it in our database.

Reference

- Zendesk (2015, Aug. 20). Introducing Maxwell, a mysql-to-kafka binlog processor. Planet MySQL. <https://planet.mysql.com/entry/?id=5992024>
- <https://www.maddyness.com/uk/2021/12/02/from-gamestop-to-tesla-how-social-media-is-driving-the-stock-market/>
- [Ghulam, H., Zeng, F., Li, W., & Xiao, Y. \(2019\). Deep learning-based sentiment analysis for roman urdu text. Procedia computer science, 147, 131-135.](https://www.sciencedirect.com/science/article/pii/S1877050919300017)
- Rao, T., & Srivastava, S. (2012). Analyzing stock market movements using twitter sentiment analysis.

Thank you