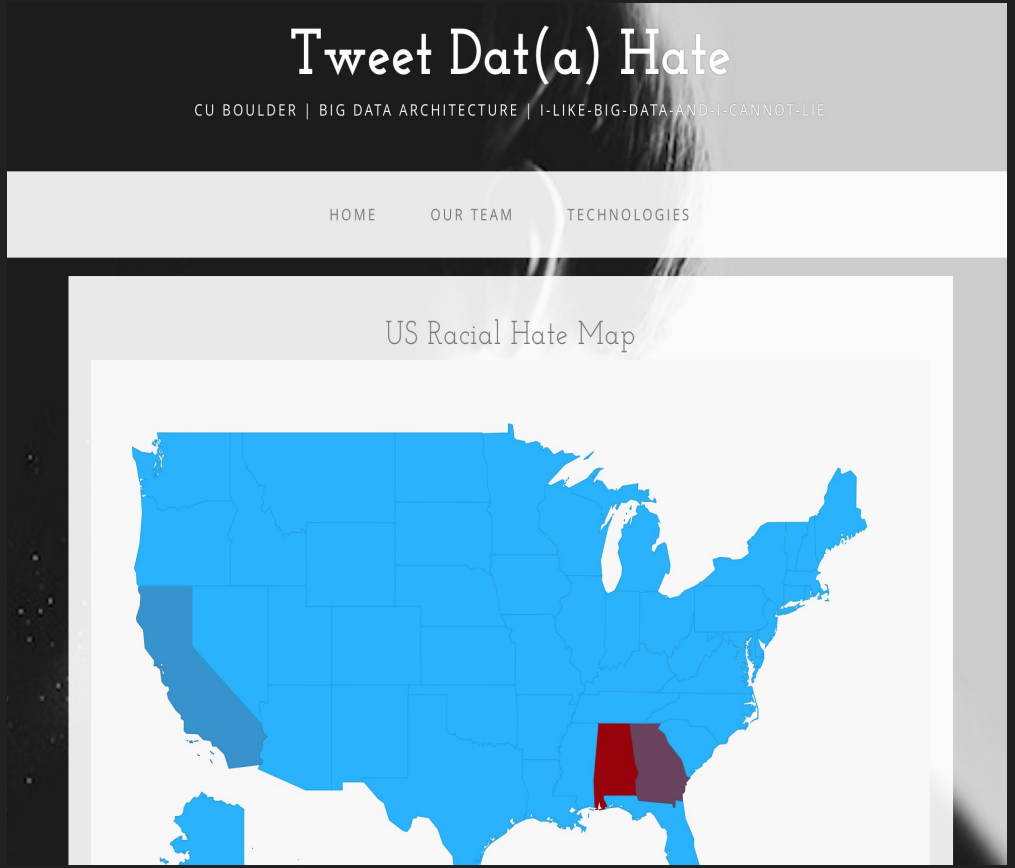
Tweet Dat(a) Hate

Team: I-Like-Big-Data-And-I-Cannot-Lie

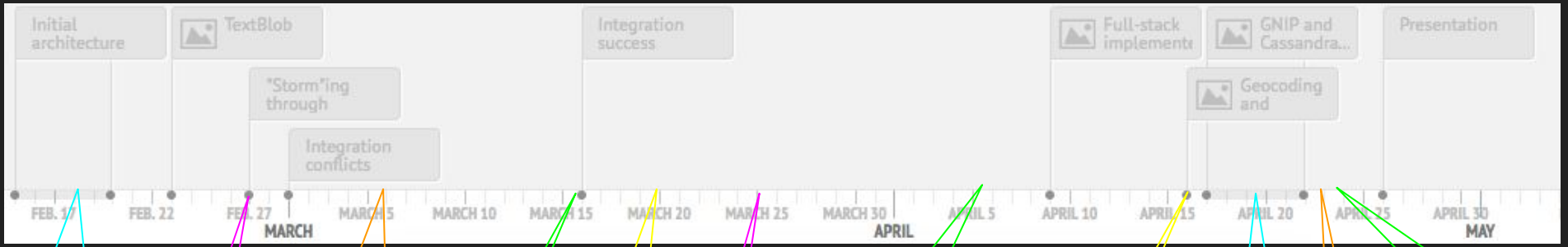
Amber Womack, Christine Nguyen, Erik Eakins
Joel Marquez, Michael Swisher

Story of Project

- Analyze % of hate/racism tweets relative to total number per state
- Full-Stack streaming data from Twitter API/GNIP
- Why it mattered to us?
  - Twitter allows people to express opinions/feelings
  - Visualize recent political sentiments and societal changes
  - Region most prevalent for racism



Execution Timeline



Storm instance set-up

Cassandra Setup

Storm intakes Twitter streaming API

Sentiment Analysis works with Storm

Initial Elm frontend

Storm deployed to server

Storm connects to Cassandra & AWS servers up

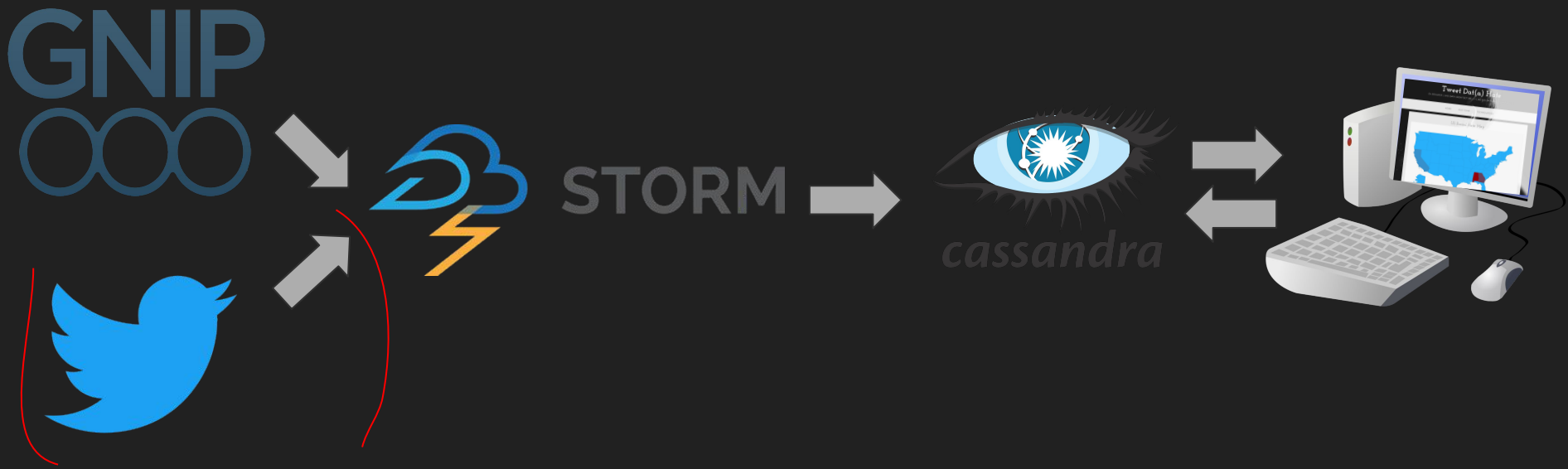
Attempted Elassandra

Elassandra breaks Cassandra

Cassandra Refactor - Storm works with GNIP

Switch from Twitter to GNIP

Architecture



Storm

- Multi-language (Java, Python)
- Spout consumes GNIP Search API
  - 500 tweet objects/request
  - Broken down into single tweet objects & queued
  - Continuously makes GNIP requests using “next” token
- Formatter bolt
  - Scrapes only the info we require from tweet object
  - Formats into custom object
- Sentiment Analysis bolt
- Insert bolt
- Print bolt
  - Used to log errors from python bolts

```
TopologyBuilder builder = new TopologyBuilder();

builder.setSpout("twitter", new TwitterStreamSpout(gnipUser, gnipPass, gnipUrl));
builder.setBolt("formatter", new TweetFormatterBolt()).shuffleGrouping("twitter");
builder.setBolt("sentiment", new SentimentAnalysisBolt()).shuffleGrouping("formatter");
builder.setBolt("insert", new CassandraInsertBolt()).shuffleGrouping("sentiment");
builder.setBolt("print", new TwitterStreamPrint()).shuffleGrouping("insert");
```

Sentiment Analysis

```python
from textblob import TextBlob
from textblob.sentiments import NaiveBayesAnalyzer
import math
import storm
#For this to work, it is necessary to use the following command:
#pip install textblob
#The test.txt uses hate words to say positive stuff about racism, etc
#The score will be greater than 0 because the statement is positive
#even though the content is not.


class tweetFilteredSentiment(storm.BasicBolt):
    hateWords = set()

    def normalizeText(self,tweet):
        tempString.deepcopy(tweet)
        # Do tolkenization and normalization of tweet
        return tempString

    def checkForHateWords(self,tweet):
        #normalized = normalizeText(tweet)
        for word in tweet.split():
            if word.lower() in hateWords:
                sentimentScore = TextBlob(tweet)
                storm.emit((True,abs(sentimentScore.sentiment.polarity)))
        storm.emit((False,0))
```

```python
    def loadHateWords(self):
        hateWords = set()
        hateFile = open("hate.txt", "r")
        for word in hateFile:
            hateWords.add(word.rstrip())
        hateFile.close()

#tweetFilteredSentiment().run()

# The main function will need to be removed when
# integrating with Storm's java program.
def main():
    loadHateWords()
    testSentiment = open("test.txt","r")
    for line in testSentiment:
        result,score = checkForHateWords(line)
        if result:
            print("{} score:{}".format(line.rstrip(),score))
    testSentiment.close()

main()
```

Cassandra

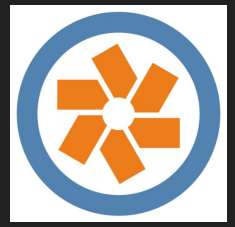
- Originally 3.0.12 and moved to Elassandra
- Deployed on new server, but ES failure
- Real Data changed to GNIP
- Python bolt will insert Tweet data to Cassandra

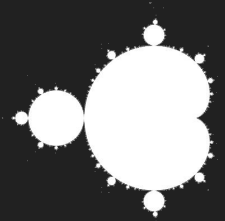
Elm

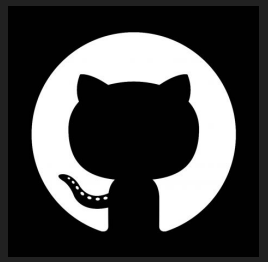
```
import Bootstrap.CDN as CDN
import Bootstrap.Navbar exposing (..)
import Bootstrap.Grid as Grid
import Bootstrap.Grid.Col as Col
import Bootstrap.Grid.Row as Row
import Html exposing (..)
import Html.Attributes exposing (..)
import Html.Events exposing (..)

main =
    Html.program
        { init = init
        , view = view
        , update = update
        , subscriptions = subscriptions
        }

-- MODEL
type alias Model =
    { tech: String
    , image: String
    }


-- UPDATE
type Msg =
    Title
    | Header String

update: Msg -> Model -> (Model, Cmd Msg)
update msg model =
    case msg of
        Title ->
            (Model "Tweet Dat(a) Hate" "", Cmd.none)
        Header newHeader ->
            (Model newHeader "", Cmd.none)


-- VIEW
view: model -> Html Msg
view model =
    Grid.container
        [style
            [ ("padding-left", "15px")
            , ("padding-right", "45px")
            , ("margin-left", "15px")
```

- Used for Tech page
- Extra for learning technologies
- Bootstrap library

Other Technologies Used







100%

pig with lipstick only wishes…..

Project Site: http://swishertest.site