

## Project: Twitter Stream Collector and Reporting System (TWCR)

GIT :

TW\_Collector : [https://github.com/afigueroa14/tw\\_collector.git](https://github.com/afigueroa14/tw_collector.git)TW\_REST : [https://github.com/afigueroa14/tw\\_rest.git](https://github.com/afigueroa14/tw_rest.git)

## Summary

The TWCR System is designed base on distributed architecture (Figure 1). There are two applications that compose the TWCR system. The first application (Twitter Collector) collects the Twitter Data, and second application (Twitter Rest) allows the user to query for the information. Both applications are connected using the Apache Ignite, which allows data to pass between different nodes on the Application Cluster. The Twitter Collector uses the SPARK Big Data Engine to manage the high volume of Twitter Stream.

## Application Architecture

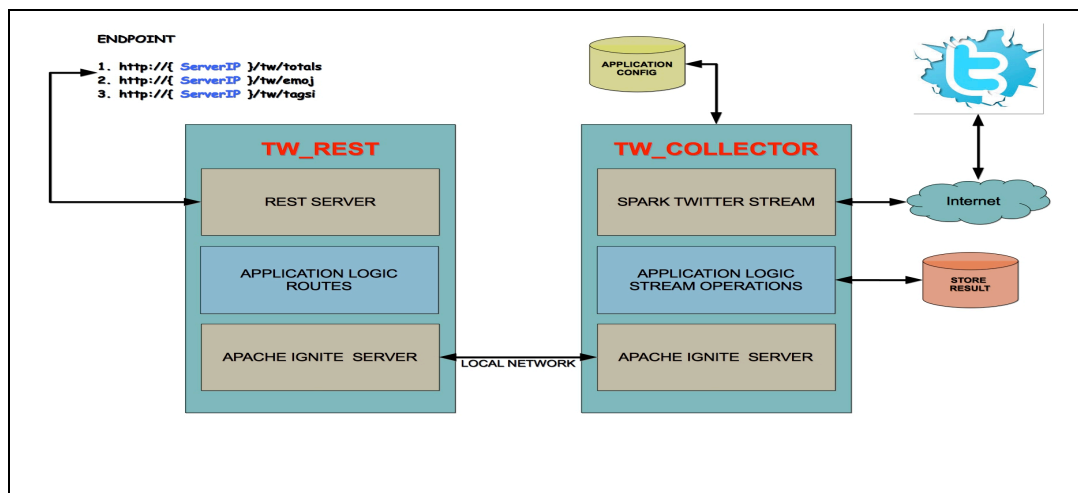


Figure 1. Twitter Architecture

TW-REST EndPoints that allows the user accesses the Information.

Requirements	Rest EndPoint – For Full Request	Condition
Version	<a href="http://{ ServerIP }/tw/version">http://{ ServerIP }/tw/version</a>	
Total Number of Tweets Received	<a href="http://{ ServerIP }/tw/totals">http://{ ServerIP }/tw/totals</a>	
Average Tweets per hour/minute/second	<a href="http://{ ServerIP }/tw/totals">http://{ ServerIP }/tw/totals</a>	
Top emojis in Tweets	<a href="http://{ ServerIP }/tw/emoji">http://{ ServerIP }/tw/emoji</a>	?maxvalue=###
Percent of Tweets that Contains Emojis	<a href="http://{ ServerIP }/tw/totals">http://{ ServerIP }/tw/totals</a>	
Top HashTags	<a href="http://{ ServerIP }/tw/hashtags">http://{ ServerIP }/tw/hashtags</a>	?maxvalue=###
Percent of Tweets that Contain a url	<a href="http://{ ServerIP }/tw/totals">http://{ ServerIP }/tw/totals</a>	
Percent of Tweets that contain a photo url	<a href="http://{ ServerIP }/tw/photo">http://{ ServerIP }/tw/photo</a>	?maxvalue=###
Top Domains of Url in Tweets	<a href="http://{ ServerIP }/tw/domaintop">http://{ ServerIP }/tw/domaintop</a>	?maxvalue=###
URL on Tweets	<a href="http://{ ServerIP }/tw/urltop">http://{ ServerIP }/tw/urltop</a>	?maxvalue=###