

Курсовой проект

Тема: Разработка пользовательской активности на streams meetup с распределением по странам

+

o

•



Выполнил:
Кузнецов Виктор Владимирович

Проект разработан для решения следующей бизнес-цели:

Предоставление информационной панели по пользовательской активности в Streams MeetUp в группах расположенных в разных странах.

Для решения данной задачи было принято решение:

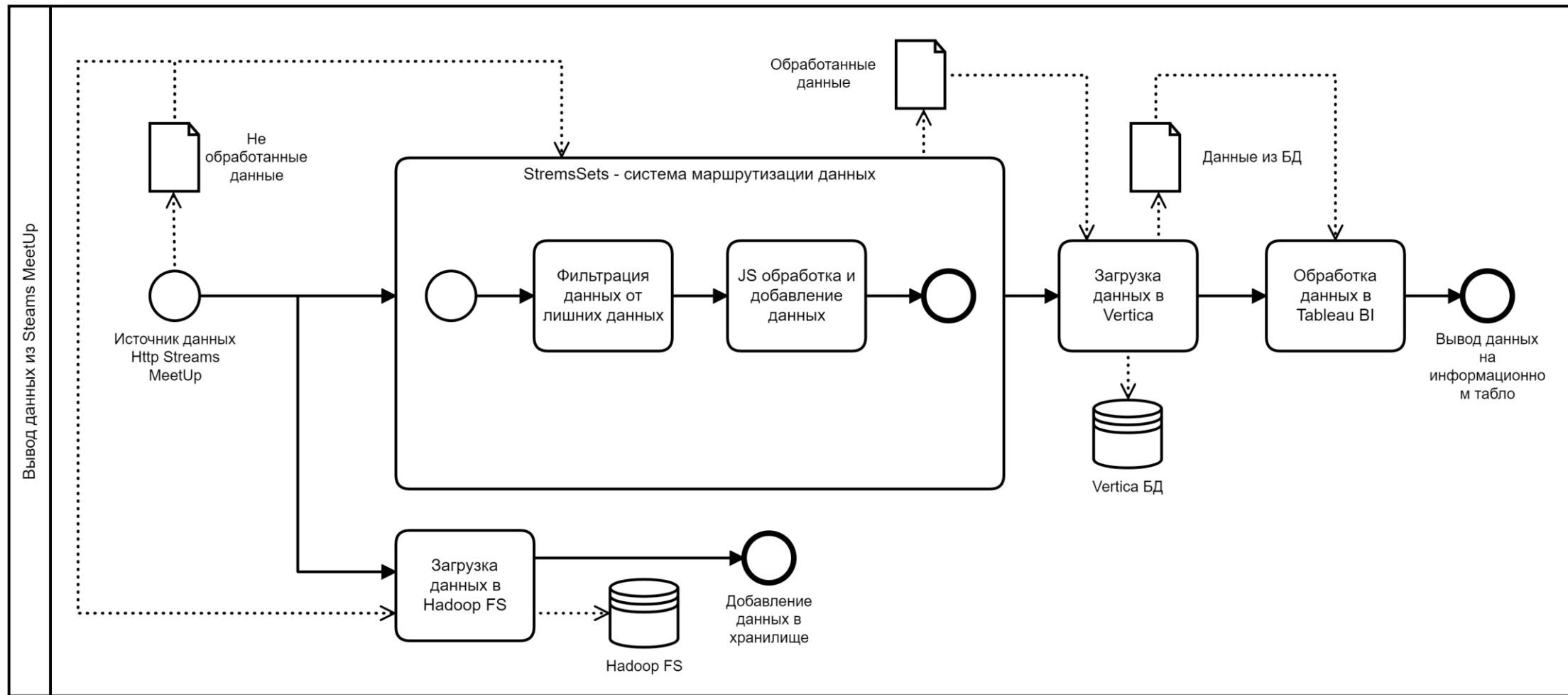
1. Сбор информации с ресурса Streams MeetUp
2. Резервирования данных в Data Lake (Hadoop FS)
3. Трансформация данных и загрузка в аналитическую базу данных Vertica
4. Предоставление информационной панели Tableau по пользовательской активности в Streams MeetUp в группах расположенных в разных странах.

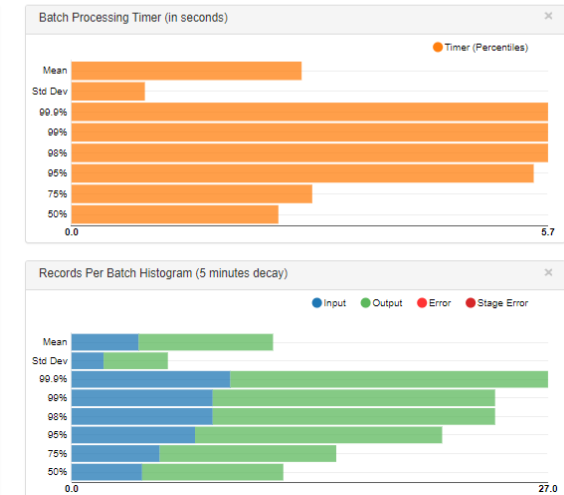
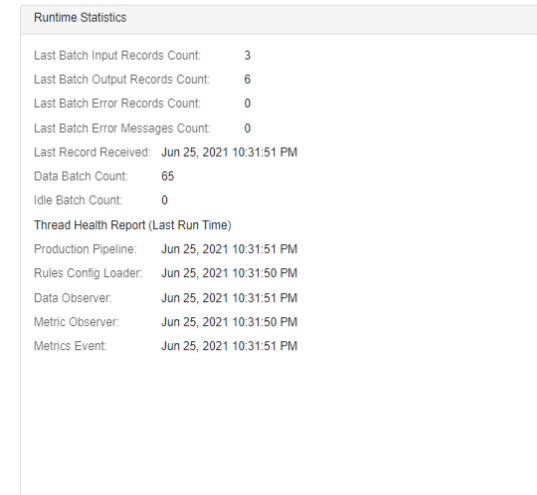
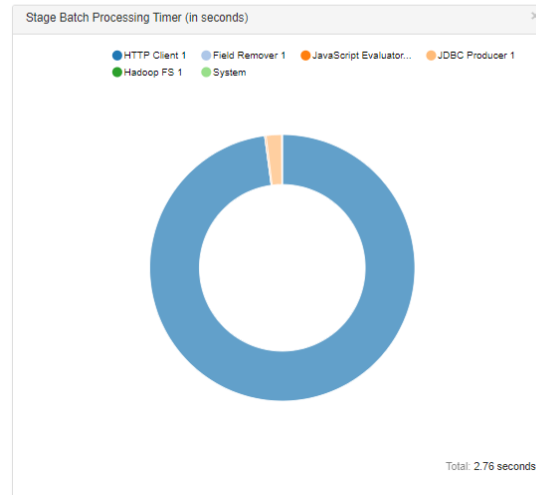
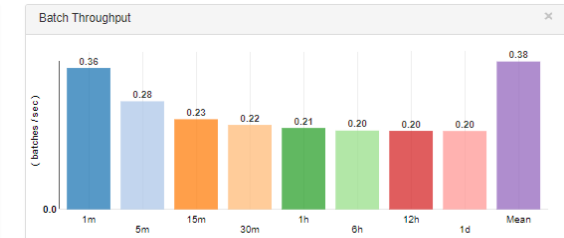
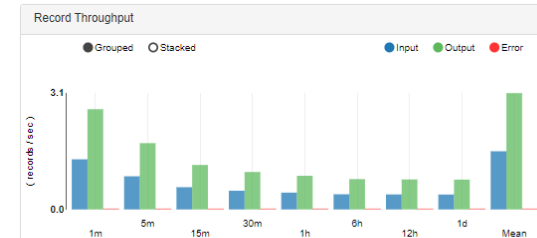
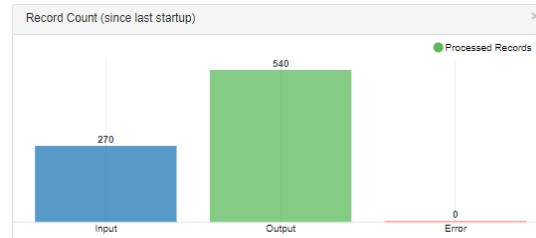
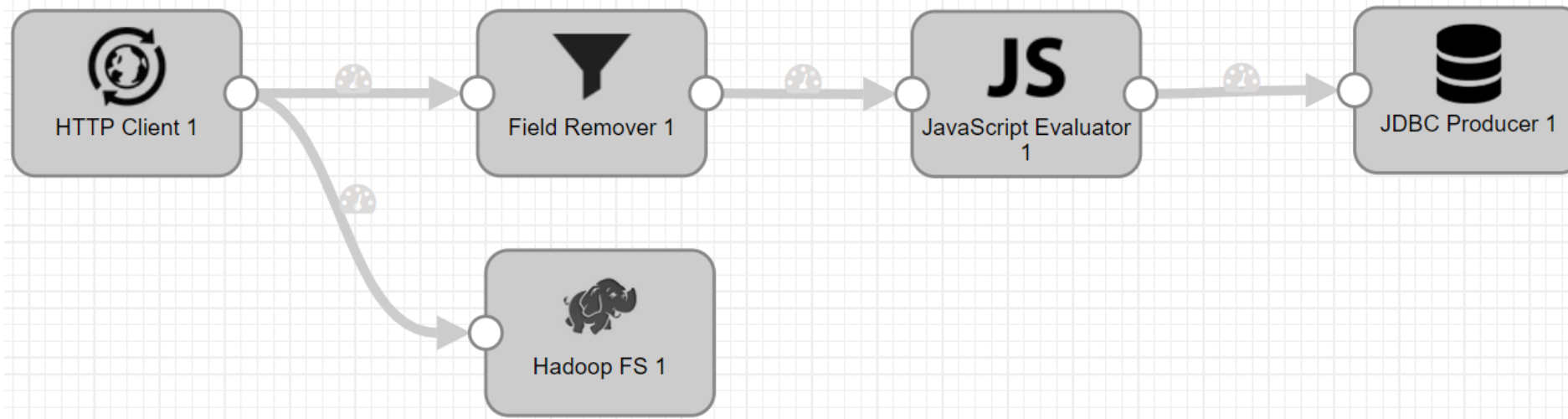
Для выполнения данных требований предусмотрен следующий состав программного обеспечения:

- StreamsSet
- Hadoop FS
- Vertica VM
- Tableau

Данные технологии позволяют в минимальные сроки реализовать данный проект.

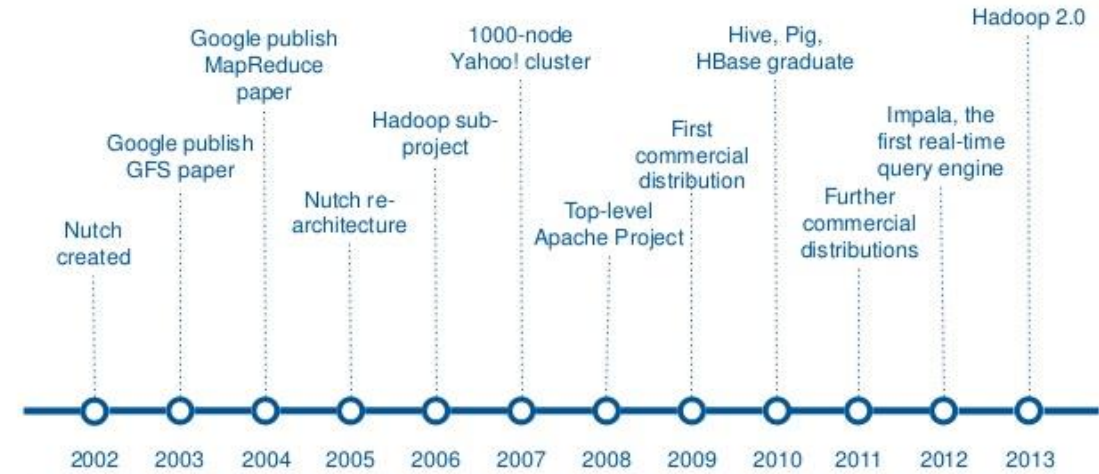
Архитектура проекта



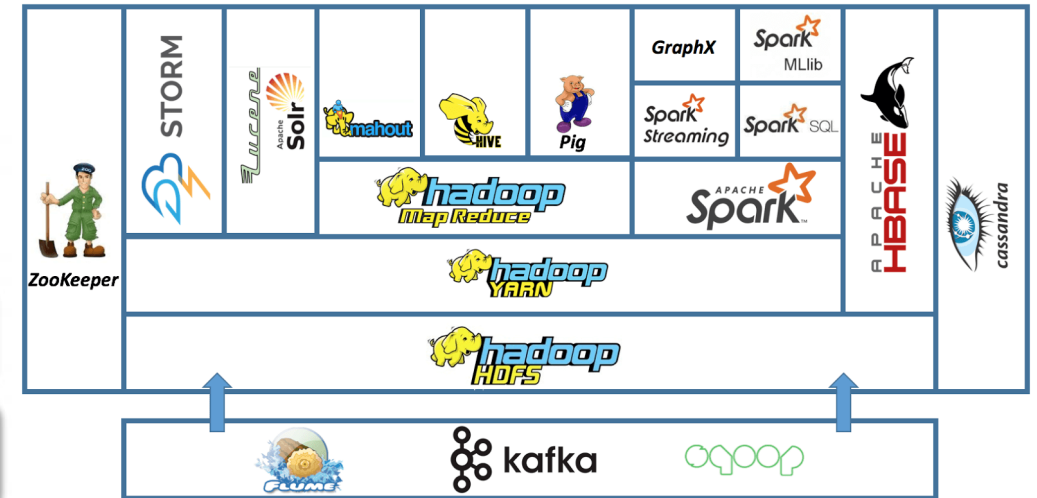
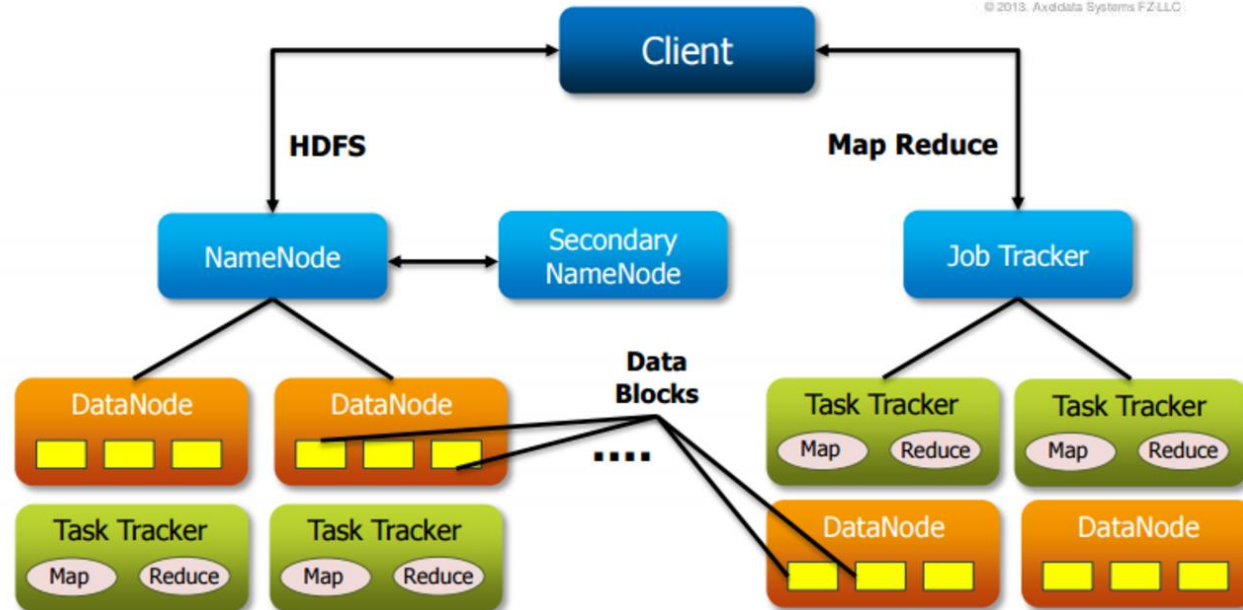




A Brief History of Hadoop



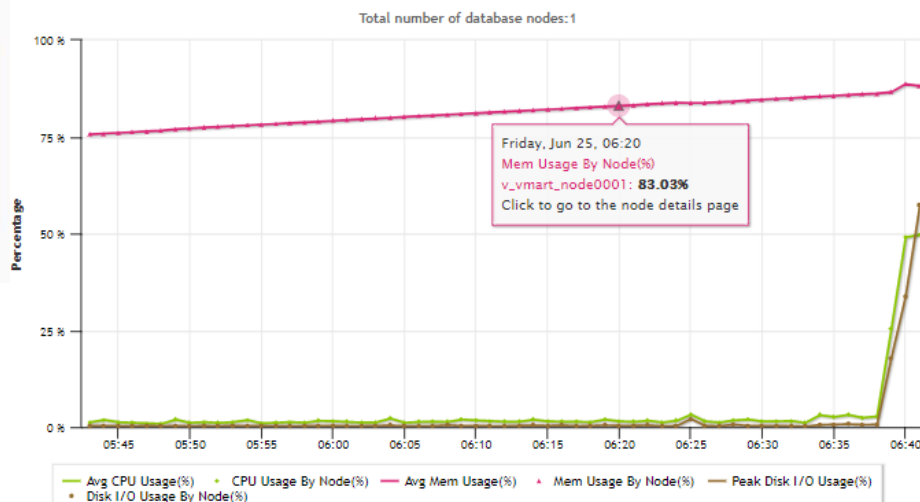
© 2013, Axidata Systems FZ LLC



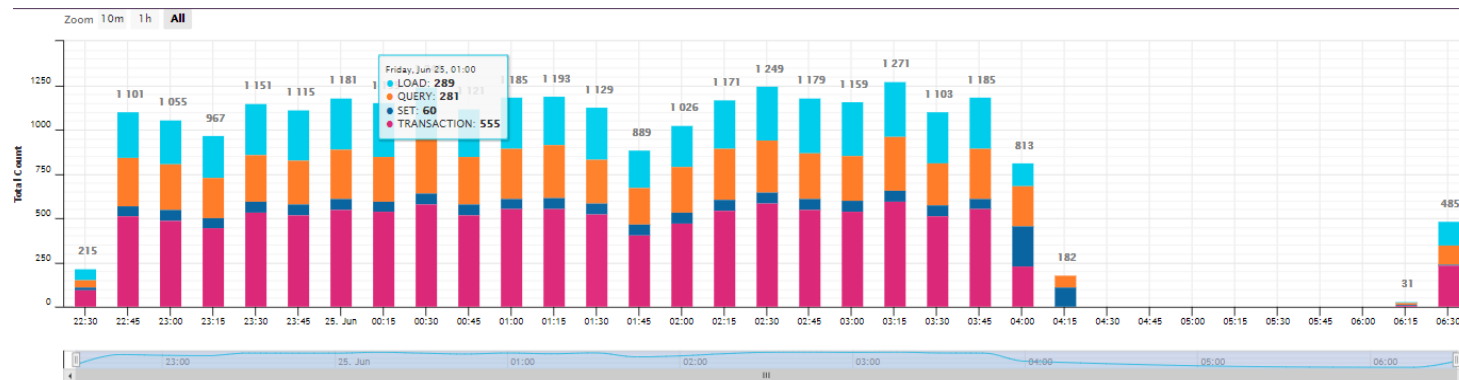
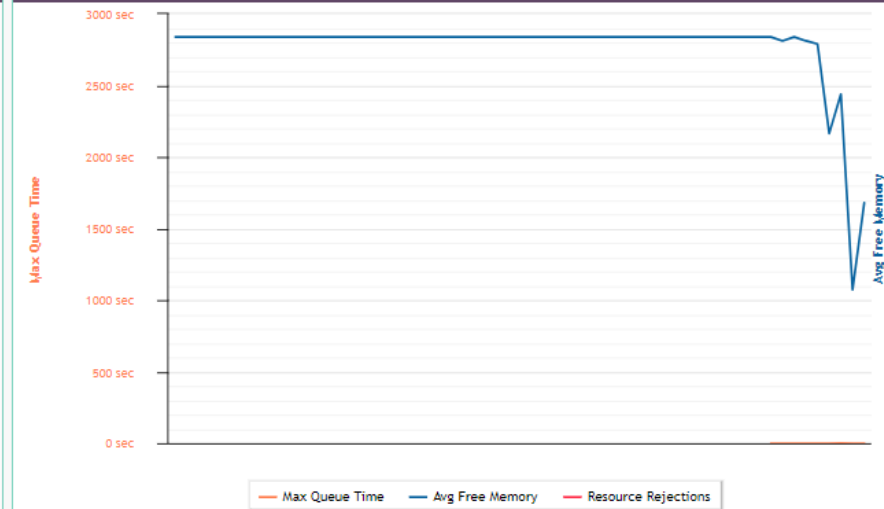


IP: 127.0.0.1
 enp0s17: 10.0.2.15
 virbr0: 192.168.122.1
 Name: v_vmart_node0001
 Status: UP
 Total mem: 3791 MB
 CPU count: 2
 Health Score: 1/100

PU/Memory/Disk I/O (Showing 58 minutes)



Database General Pool Usage (Showing 59 minutes)

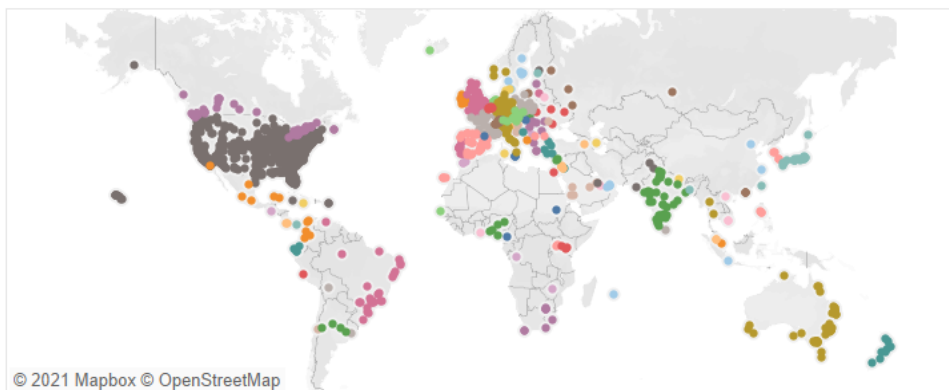


```
CREATE TABLE stream_meetup(
  stream_meetup_id IDENTITY(1,1) NOT NULL PRIMARY KEY,
  guests INTEGER,
  rsvp_id INTEGER,
  mtime DECIMAL(18,2),
  date_create VARCHAR(32),
  is_read INTEGER,
  venue_id INTEGER,
  venue_name VARCHAR(256),
  venue_lon DOUBLE PRECISION,
  venue_lat DOUBLE PRECISION,
  member_id INTEGER,
  member_name VARCHAR(256),
  member_photo VARCHAR(256),
  event_id VARCHAR(128),
  event_name VARCHAR(256),
  event_time DECIMAL(18,2),
  event_url VARCHAR(256),
  group_id INTEGER,
  group_city VARCHAR(256),
  group_country VARCHAR(64),
  group_name VARCHAR(256),
  group_urlname VARCHAR(256),
  group_lon DOUBLE PRECISION,
  group_lat DOUBLE PRECISION,
  topics VARCHAR(2048)
);
```

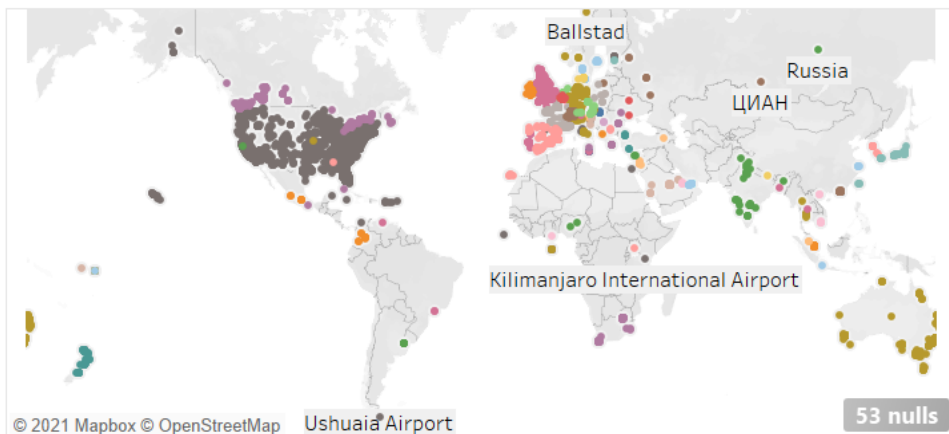


Распределение активности для StreamsMeetUp

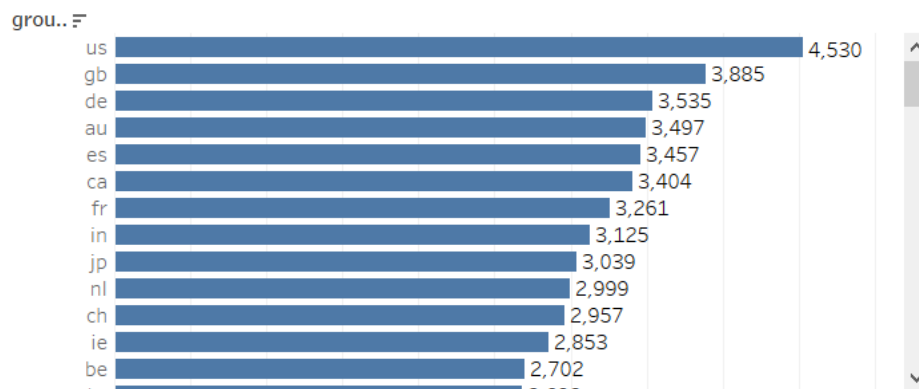
Группы по странам



События по странам



Количество встреч по странам



Количество встреч по часам для каждой станы

