Разработка системы сбора, анализа, хранения и представления журналов событий сетевых служб

В рамках конкурса «Открытый регион. Хакатон-2016» Кейс ООО «ИТ-Групп»

Докладчик: Гайнанов Руслан

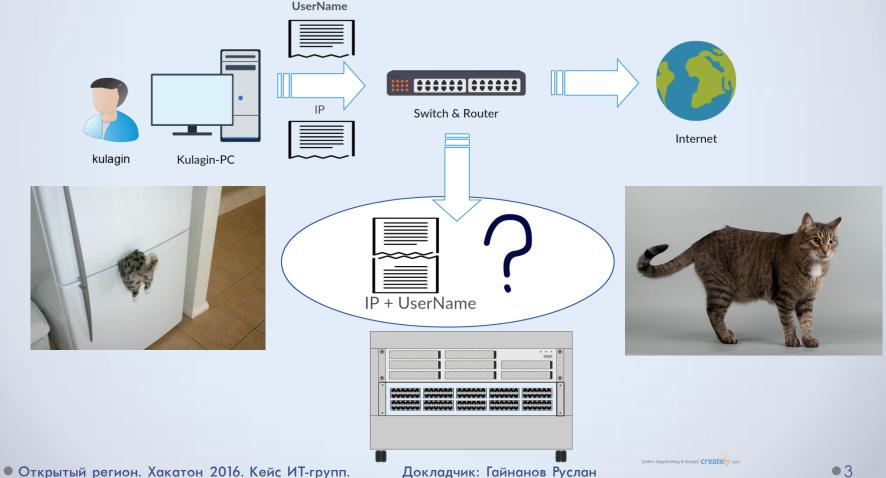
Проблема

- Кто, когда и с какого устройства работал в сети с заданным IP адресом
 - ⇒ чтобы знать сетевую активность пользователя (время работы, посещаемые сайты, потребляемый трафик и т.д.)
 - ⇒чтобы знать местоположение пользователя
 - ⇒ чтобы лучше управлять сетевыми службами



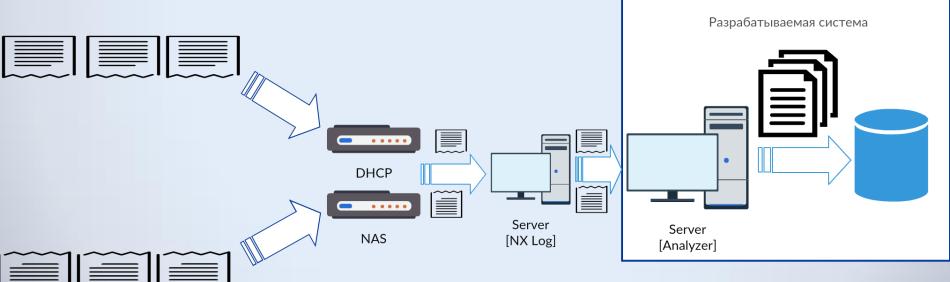
Проблема

• Информация о сетевом соединении (IP-адрес, маска и пр.) и информация о пользователе (UserName, групповые политики и пр.) формируются разными независимыми службами



Задачи

- Разработать систему сбора и хранения событий от DHCP и NAS серверов получаемых от NX Log по UDP-соединению
- Разработать систему анализа данных для корреляции событий
- Разработать систему визуализации хранимых событий



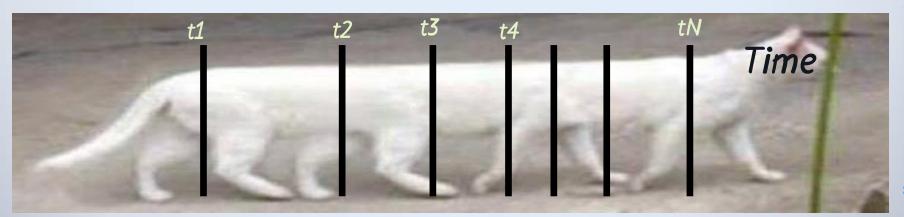
Решение

- Сбор и анализ данных:
 - о программа на Ruby получает данные в JSON формате по UDP, производит анализ и запись их в БД.
- Хранение данных:
 - InfluxDB база данных для хранения временных рядов (time-series databases, TSDB) – наиболее популярная за 2016 согласно агентству DB Engines.
- Визуализация данных:
 - Grafana удобный веб-сервис для визуализации всевозможных метрик и работы с TSDB.



InfluxDB

- Open-source DBMS. Написана на Go
- В основе БД механизм Time-Structured Merge Tree (TSM)
- Отсутствуют внешние зависимости
- Поддерживает кластеризацию
- Запросы на SQL-like по HTTP API
- Продолжительность хранения данных обеспечивается с помощью политик



Демонстрация работы http://159.93.36.108:3000

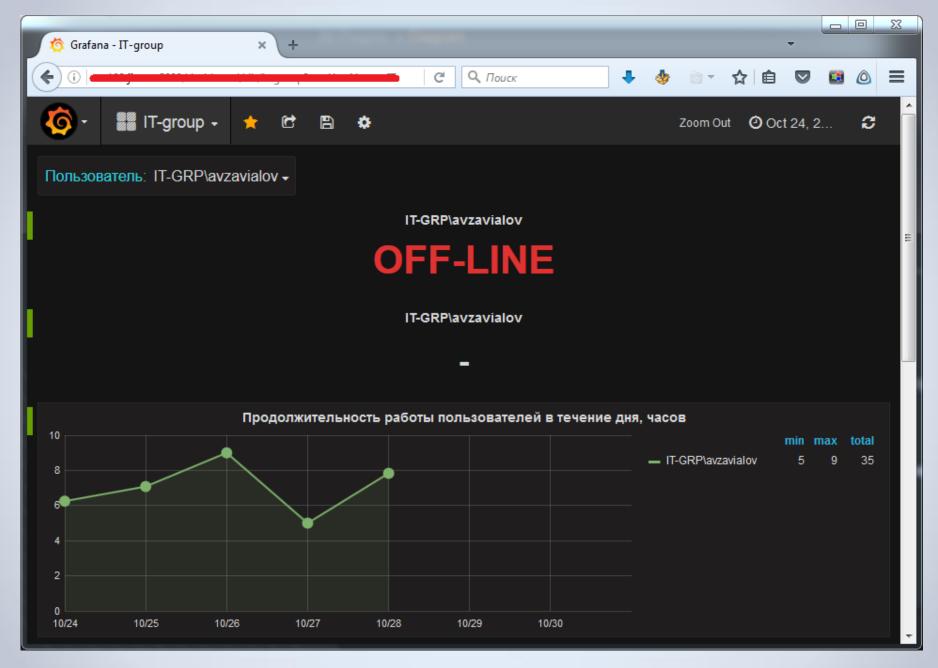
(vm108.jinr.ru:3000)

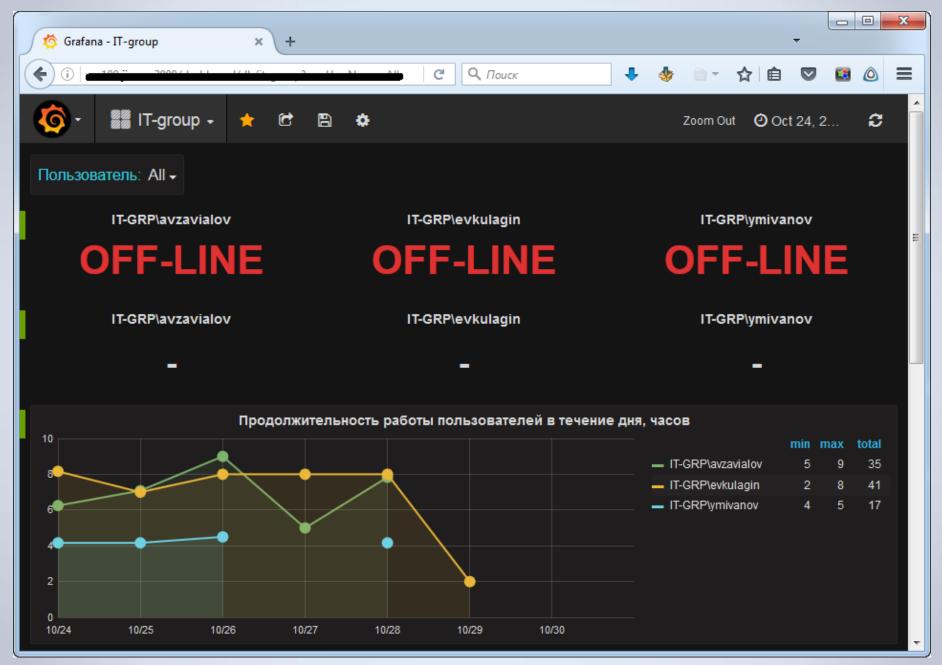
user: itgrp

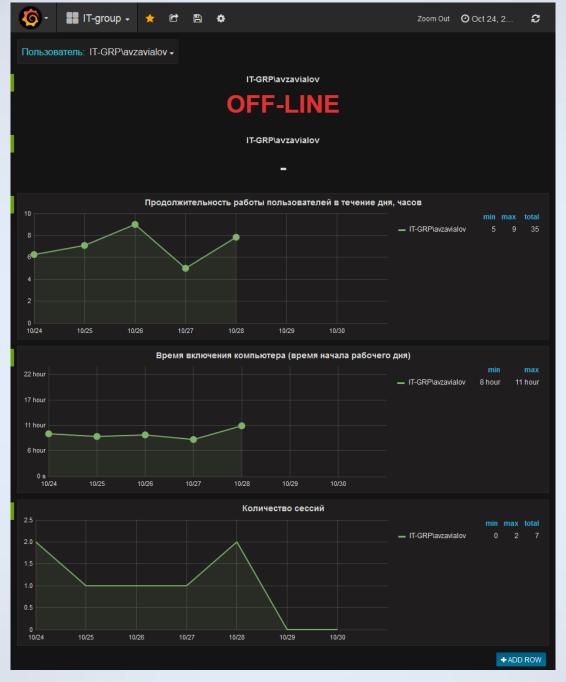
password: hackaton2016

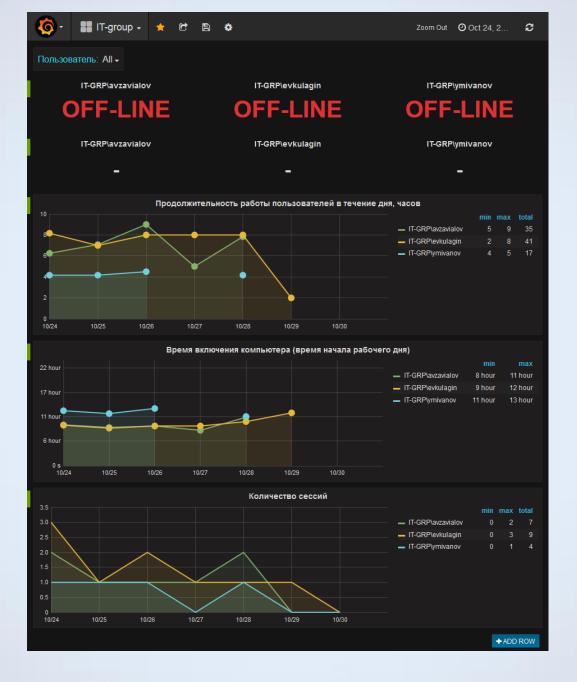


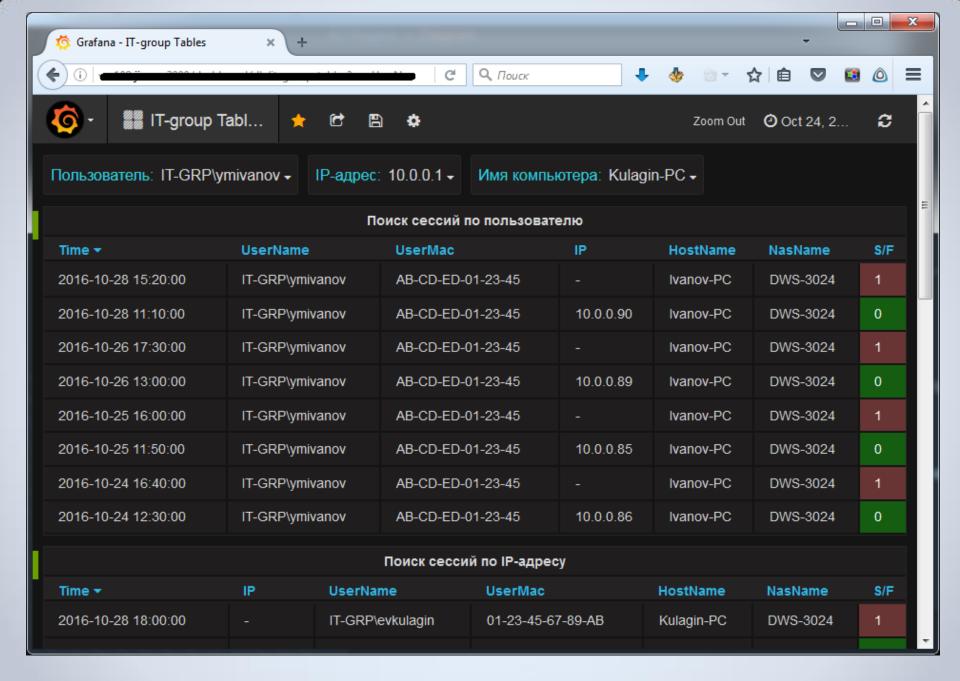
В конец (19)

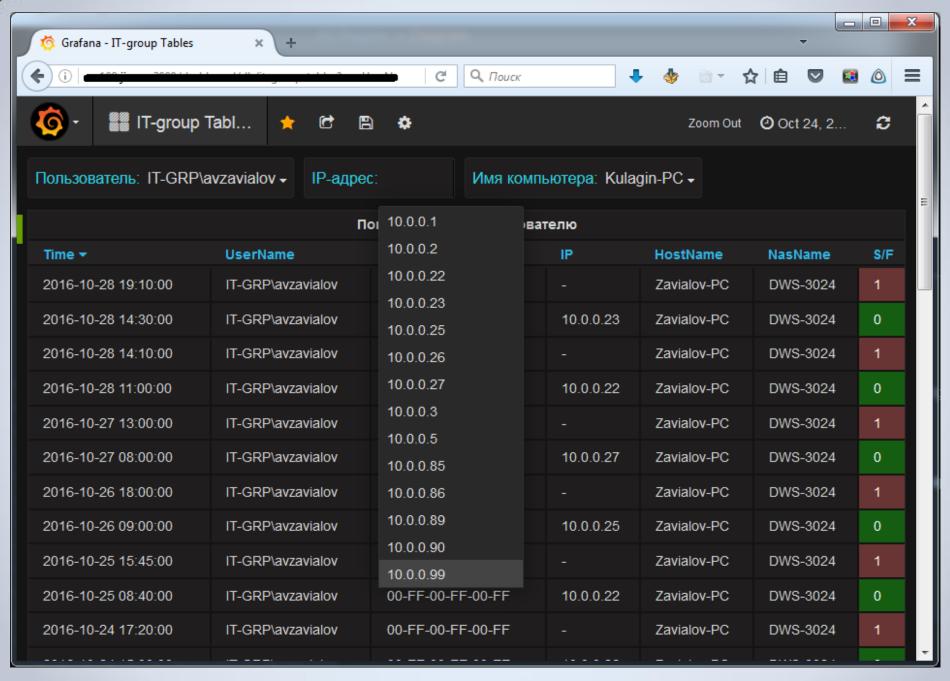




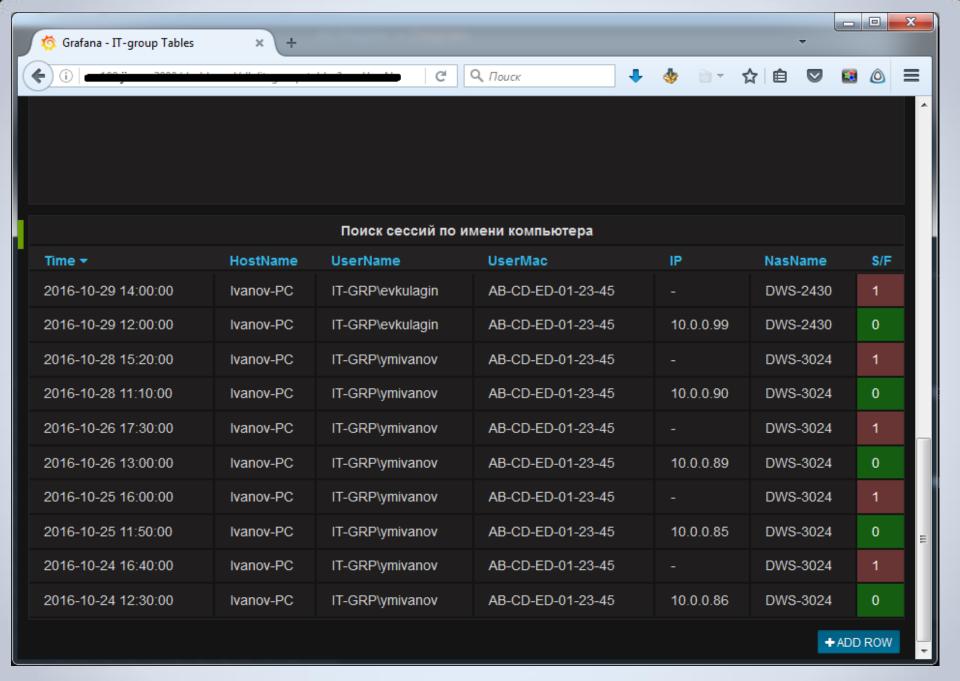


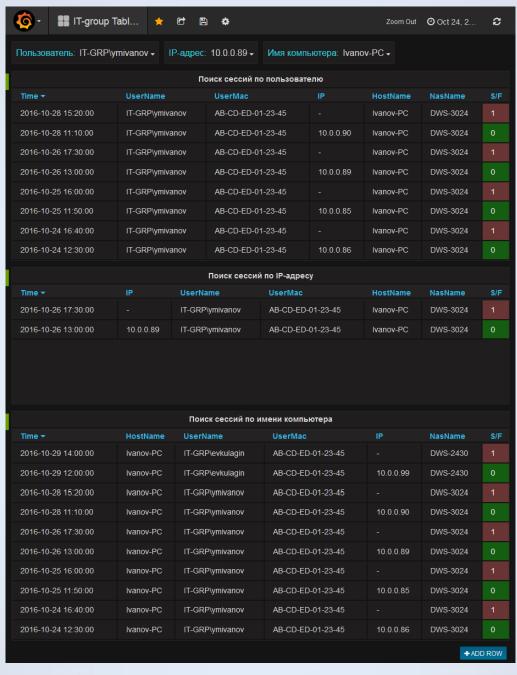




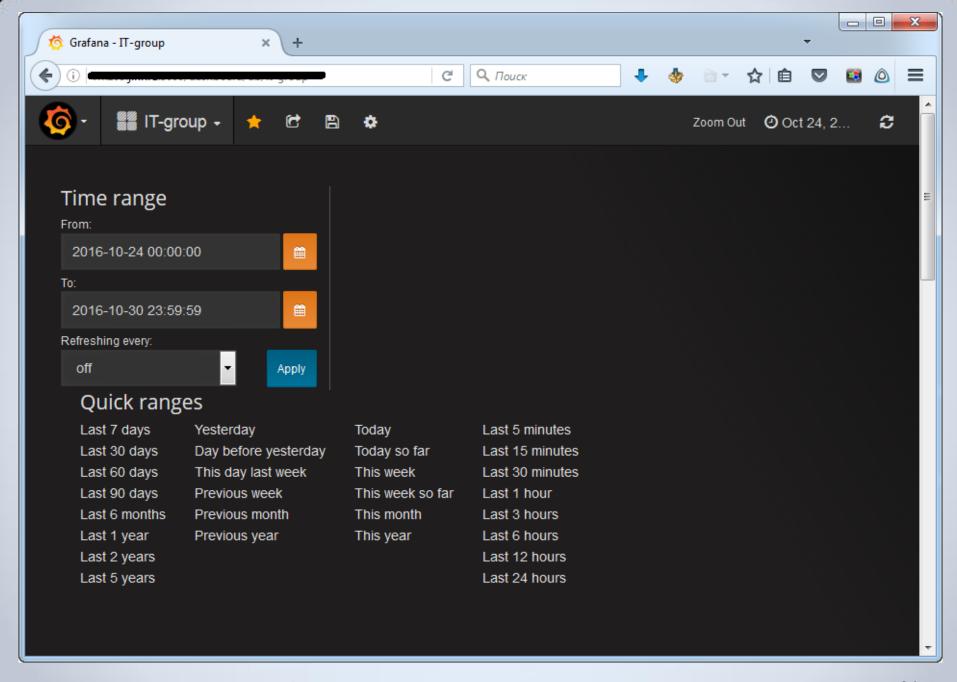


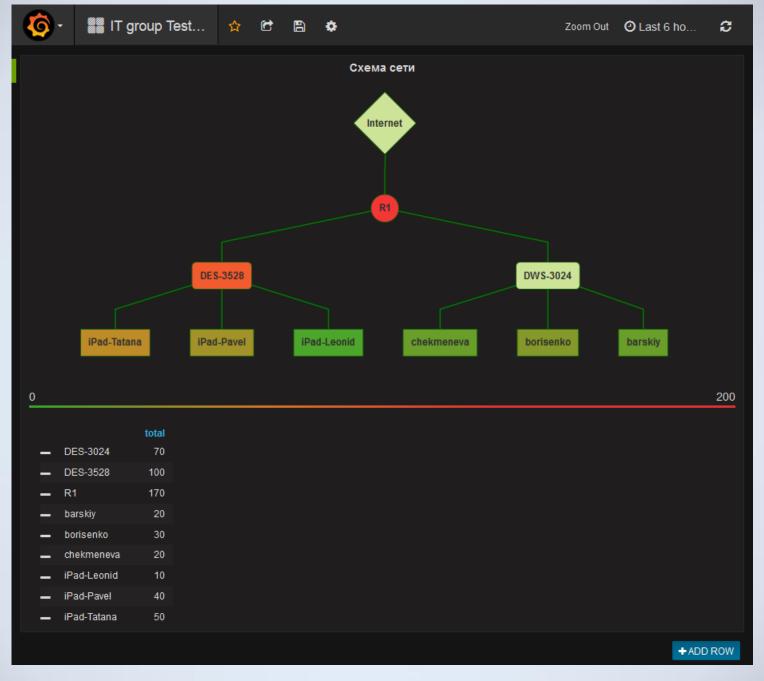
[•] Открытый регион. Хакатон 2016. Кейс ИТ-групп.

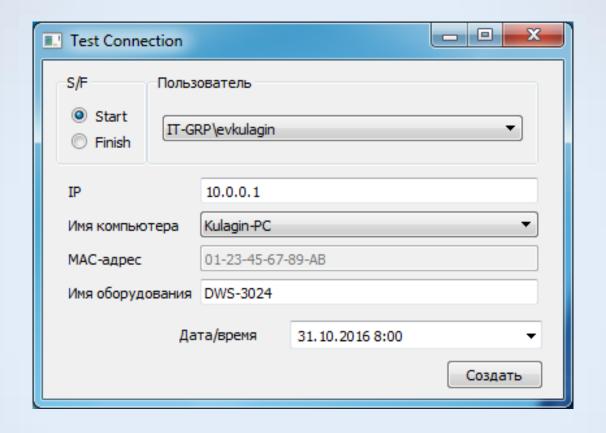


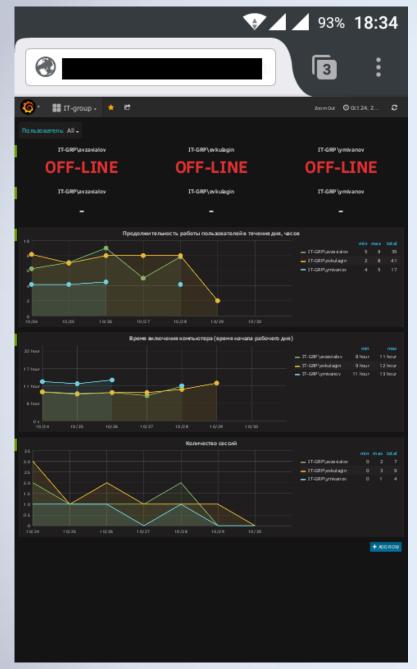


[•] Открытый регион. Хакатон 2016. Кейс ИТ-групп.











HTTP POST-запросы [1]

Узнать какой(-ue) IP-адрес(а) были выданы пользователю с именем kulagin 25/10/2016

```
[root@vm-hackathon-2 ~]# curl -GET
q=SELECT z IP FROM connections WHERE time >= '2016-10-25' AND time<'2016-
10-26' AND UserName=~/.*kulagin.*/"
   "results": [
            "series": [
                    "name": "connections",
                    "columns": [
                        "time",
                        "z IP"
                    "values": [
                            "2016-10-25T03:30:00Z",
                            "10.0.0.5"
                            "2016-10-25T10:30:00Z",
```

HTTP POST-запросы [2]

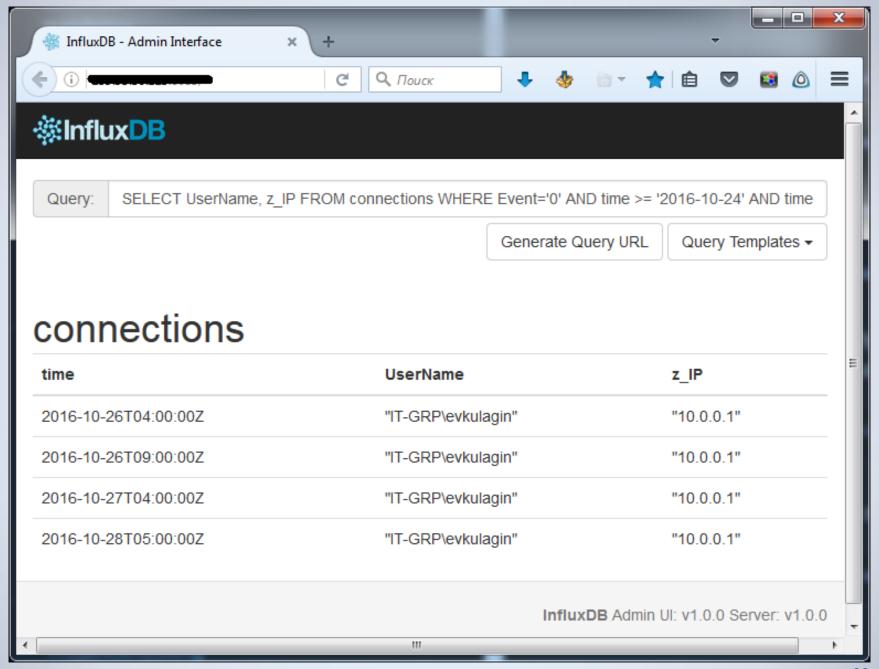
Получить суммарное время работы пользователей принадлежащих домену IT-GRP за выбранный период

```
[root@vm-hackathon-2 ~]# curl -GET
http://localhost:8086/query?pretty=true&db=telegraf' --data-urlencode
'q=SELECT sum(z D)/3600 FROM connections WHERE time >= '2016-10-24' AND
time < '2016-10-30' AND UserName =~ /IT-GRP.*/"
    "results": [
            "series": [
                    "name": "connections",
                    "columns": [
                        "time",
                        "sum"
                    "values": [
                            "2016-10-24T00:00:00Z",
                            219.41666666666666
```

HTTP POST-запросы [3]

Узнать, кто заходил под IP-адресом 10.0.0.1 в течение заданного времени

```
root@vm-hackathon-2 ~]# cur
 SELECT UserName, z IP FROM connections WHERE Event='0' AND time >=
 016-10-24' AND time < '2016-10-30' AND Ip='10.0.0.1'"
  "results": [
           "series": [
                   "name": "connections",
                   "columns": [
                       "time",
                       "UserName",
                   "values": [
                           "2016-10-26T04:00:00Z",
                           "IT-GRP\\evkulagin",
                           "10.0.0.1"
                           "2016-10-26T08:00:00Z",
                           "IT-GRP\\evkulagin",
                           "2016-10-26T09:00:00Z",
                           "IT-GRP\\evkulagin",
                           "2016-10-26T13:00:00Z",
                           "IT-GRP\\evkulagin",
                           "2016-10-27T04:00:00Z",
                           "IT-GRP\\evkulagin",
                           "10.0.0.1"
                           "2016-10-27T12:00:00Z",
                           "IT-GRP\\evkulagin",
```



Результаты

- Выбрана система хранения данных InfluxDB
- Написаны программы для сбора данных
- Выбрана система для визуализации хранимых данных Grafana
- Созданы панели для показа различной информации, графиков и таблиц на основе хранимых данных:
 - Текущем статусе пользователя(-ей)
 - о Текущем ІР-адресе
 - о Продолжительности сессии, общем количестве сессий в течение дня
 - История выданных IP пользователям, подключенных устройствах и пользователях
- Создана программа для генерации тестовых данных
- Представлен вариант отображения данных о потоке
- Написаны программы для анализа входящих пакетов от DHCP и NAS
 - Но имеющихся вычислительных ресурсов для их работы оказалось недостаточно



Спасибо за внимание! Вопросы?

Докладчик:

Гайнанов Руслан Рамилевич,

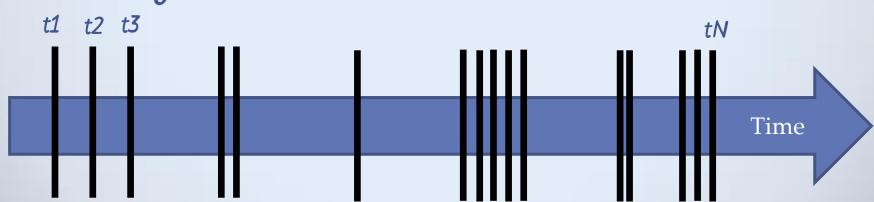
магистрант каф. ИТАС ПНИПУ

email: ruslan.r.gainanov@gmail.ru

тел.: 8-904-84-20-603



- Open-source DBMS. Написана на Go
- В основе БД механизм Time-Structured Merge Tree (TSM)
- Отсутствуют внешние зависимости
- Поддерживает кластеризацию
- Запросы на SQL-like по HTTP API
- Политики хранения обеспечивают продолжительность хранения данных



InfluxDB vs. PostgreSQL

	InfluxDB	PostgreSQL
Description	DBMS for storing time series, events and metrics	Based on the object relational DBMS Postgres
Database model	Time Series DBMS	Relational DBMS
DB-Engines Ranking	5.32 (Score)	318.69 (Score)
Rank	#45 (Overall)	#5 (Overall)
	#1 (<u>Time Series DBMS</u>)	#4 (<u>Relational DBMS</u>)
Technical documentation	docs.influxdata.com/influxdb	www.postgresql.org/docs/manuals
Initial release	2013	1989
Current release	v1.0.0, September 2016	9.6.1, October 2016
Implementation language	Go	С
Data scheme	schema-free	yes
Typing	Numeric data and Strings	yes
Server-side scripts	no	user defined functions
Secondary indexes	no	yes
SQL	no	yes
Triggers	no	yes

[•] Открытый регион. Хакатон 2016. Кейс ИТ-групп.

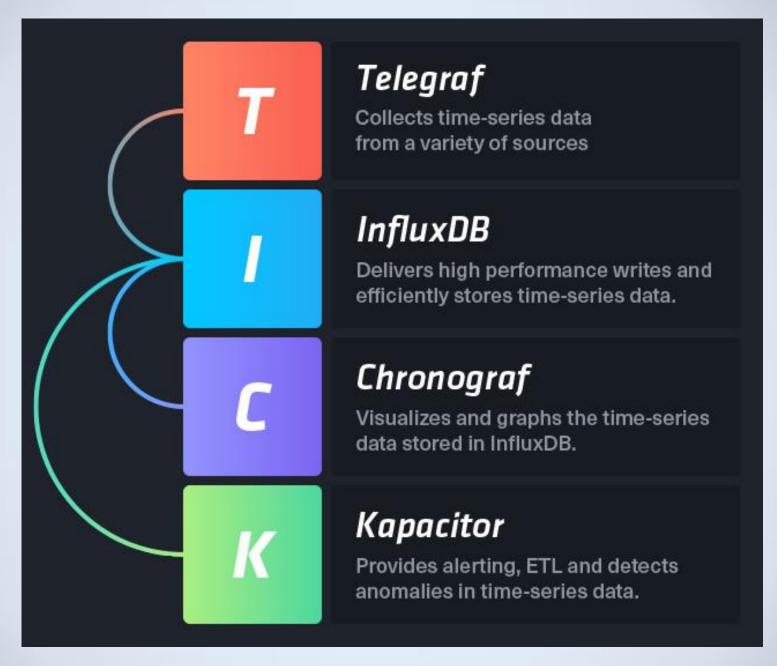
InfluxDB vs. PostgreSQL

	InfluxDB		PostgreSQL	
Server operating systems	Linux OS X		FreeBSD OpenBSD NetBSD Solaris HP-UX	OS X Linux Unix Windows
APIs and other access methods	HTTP API JSON over UDP		native C library streaming API for large objects ADO.NET JDBC ODBC	
Supported programming languages	.Net Clojure Erlang Go Haskell Java JavaScript JavaScript (Node.js)	Lisp Perl PHP Python R Ruby Rust Scala	.Net C C++ Delphi Java Perl Python Tcl	

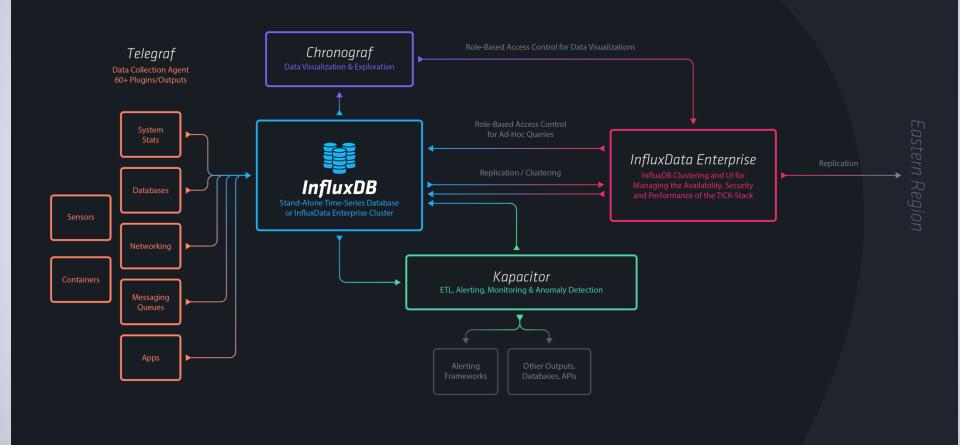
[•] Открытый регион. Хакатон 2016. Кейс ИТ-групп.







Western Region



Kapacitor

- Available event handlers:
- log log alert data to file.
- post HTTP POST data to a specified URL.
- email Send and email with alert data.
- exec Execute a command passing alert data over STDIN.
- HipChat Post alert message to HipChat room.
- Alerta Post alert message to Alerta.
- Sensu Post alert message to Sensu client.
- Slack Post alert message to Slack channel.
- OpsGenie Send alert to OpsGenie.
- VictorOps Send alert to VictorOps.
- PagerDuty Send alert to PagerDuty.
- Talk Post alert message to Talk client.
- Telegram Post alert message to Telegram client.

Kapacitor

```
stream
[alert()
      .warn(lambda: "sigma" > 2.5)
      .crit(lambda: "sigma" > 3.0)
      .log('/tmp/alerts.log')
      // Post data to custom endpoint
      .post('https://alerthandler.example.com')
      // Execute custom alert handler script
      .exec('/bin/custom_alert_handler.sh')
      // Send alerts to slack
     .slack()
      .channel('#alerts')
      // Sends alerts to PagerDuty
      .pagerDuty()
      // Send alerts to VictorOps
      .victorOps()
      .routingKey('team rocket')
```