Big Data Technology for Scientific Applications

G. SUCIU, M. ANWAR, I. ROGOJANU

A. PASAT, A. STANOIU

Outline

- Company profile
- My Profile
- Introduction
- IoT Cloud
  - Raspberry Pi
  - PyCom
  - Libelium
- Big data for academic research
- Conclusions
- Future work

Muneeb Anwar | Beia Consult International | R&D Department |ma@beia.ro

BEIA Consult

- BEIA has an experience since **1991** in over 5,000 turnkey projects for telephone and advanced communications solutions.
- Ongoing projects**: H2020 - Horizon 2020, EUREKA STARS, National Programs.**
- **ESTABLISH**, **3DSafeguard**, **WINS@HI** and **TelMonAer** are the projects which have supported our work.
- **ESTABLISH** – **Environmental Sensing To Act for a Better quality of LIfe: Smart Health**. The objective of ESTABLISH is to convert environmental (sensor) data into actionable information for users to provide a healthier and safer environment thereby improving the quality of life.
- **3DSafeguard** - proposes a solution enabling the situational awareness by introducing an integrated operation workflow.
- **WINS@HI** – Wearable IoT Network Solution for Work Safety in Hazardous Industrial Environments. In this project, we offer an ad-hoc, agile and reliable communication solution for both condition monitoring of the operations and safety of the workers in hazardous 'Industry 4.0' work environments in order to prevent unpredictable operation failures and work accidents.
- **Tel-MONAER** – Mobile System for Tele-Monitoring Air Quality. The goal of this project is to develop an IT system using the Internet of Things and Edge / Cloud Computing technologies to monitor and analyze in real time the risk factors for the environment and public health.

Partnerships

- Partners in Romanian R&D:
  - University “Politehnica” of Bucharest (www.upb.ro)
  - Romanian Space Agency (www.rosa.ro)
  - National Institute for Research and Development in Electrical Engineering (www.icpe-ca.ro)
  - National Institute of Aerospace Research “ELIE CARAFOLI” (www.incas.ro)
- Member in the Directory Council of the German-Romanian Chamber of Industry and Commerce (AHK-Deutsch-Rumaenische Industrie - und Handelskammer) and other Chambers of Commerce and Clusters
- Leader of NEM Romanian Mirror Group (www.nem-pt.ro)
- Member of Romanian Association for Electronic and Software Industry (ARIES).

Short biography

Muneeb Anwar

- Graduated from the Faculty of Electronics and Telecommunications (FILS) at the University "Politehnica" of Bucharest (UPB), Romania (www.upb.ro);

- MBAE in Business Administration at the University "Politehnica" of Bucharest (UPB), Romania (www.upb.ro);

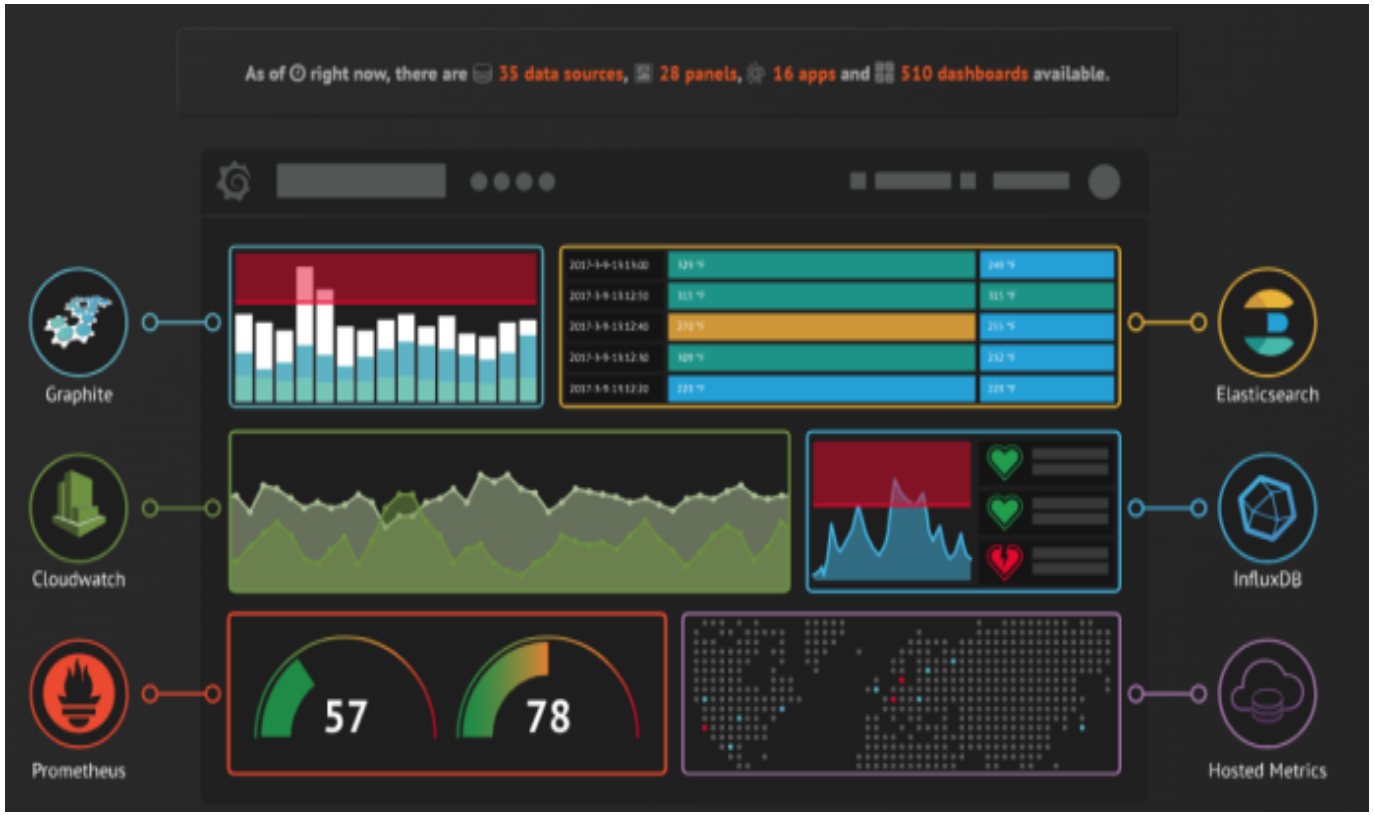
- Since 2016 Research and Development Engineer - R&D Department, working in several projects: ESTABLISH, SoMeDi, FAIR - SOLOMON + VIRTUOSE and 3DSafeguard.

Introduction

- Data processing has become a major research topic of modern science, involving several challenges related to data visualization, interaction, storage and personalization.
- In this presentation, we aim to provide a brief introduction of Big Data concepts. Also, we will show a way for efficiently storing data coming from sensors connected to different devices.

IoT Cloud

- For data visualization, we used Grafana. Each Grafana table is a query to its supported Data Source.
- For our experiments, we have used InfluxDB a real-time analysis as Data Source.
- We will show how we can deal with data collected from sensors which are connected to several devices.



Grafana's supported Data Sources

Raspberry Pi

- Raspberry Pi is a single board computer which enables the communication with several devices and users for a large variety of application. Users can create tolerably wondrous smart devices which can collect, process and display data by connecting different types of sensors to Raspberry Pi.

Raspberry Pi Model 3B+

Raspberry Pi

Graph based on data gathered from a temperature sensor

- The graph for a temperature sensor connected to Raspberry Pi.

Data History

Time	ST1.temp
2018-09-21 10:33:11	27.18
2018-09-21 10:28:11	27.59
2018-09-21 10:23:11	28.05
2018-09-21 10:18:10	28.06
2018-09-21 10:13:10	28.00
2018-09-21 10:08:10	27.99
2018-09-21 10:03:09	28.00
2018-09-21 09:58:09	27.99

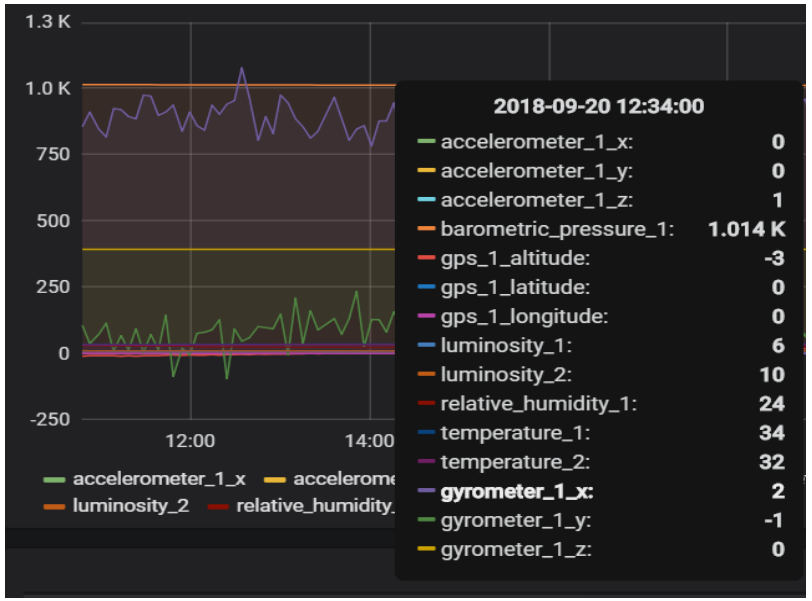
Data history for Raspberry Pi

Pycom

- PyCom PyTrack was chosen for our experiments providing sufficient coverage. Firstly, the sensors transmit data to the PyCom PyTrack board, then to the gateway and finally to the Cloud Platform.

Pycom Board

Pycom



Graph based on data gathered from Pycom's sensors

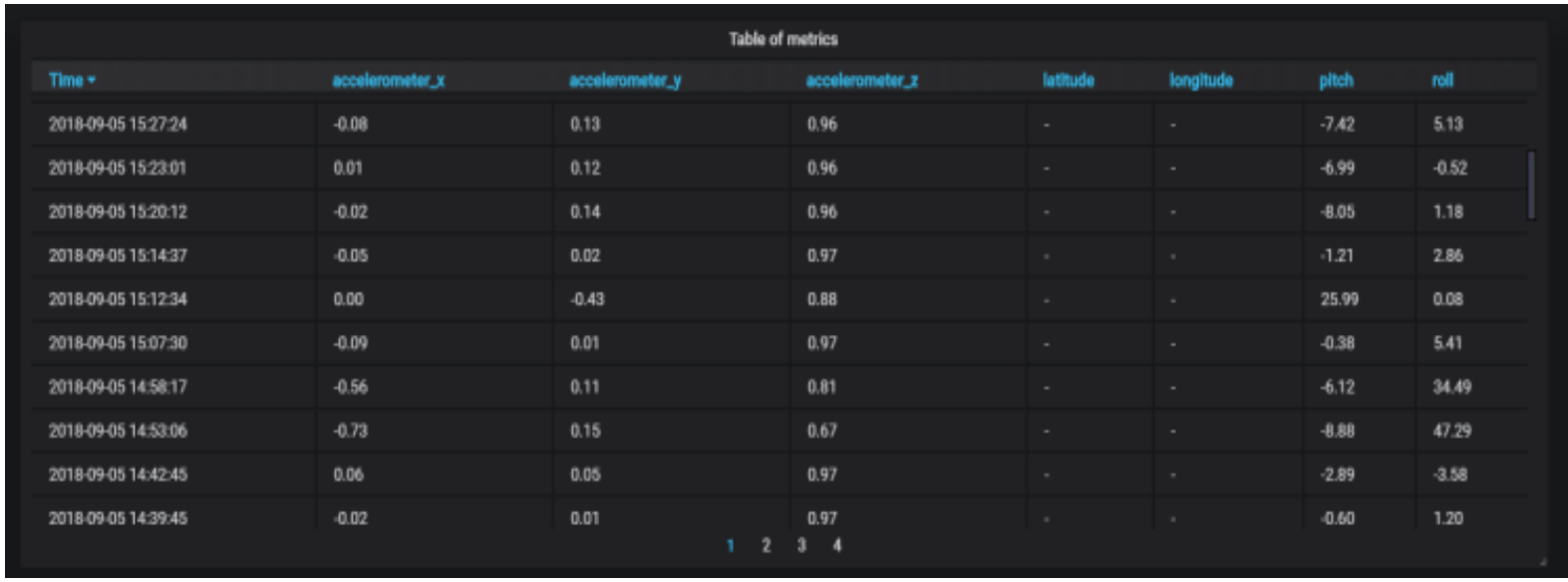


Table of metrics

Time	accelerometer_x	accelerometer_y	accelerometer_z	latitude	longitude	pitch	roll
2018-09-05 15:27:24	-0.08	0.13	0.96	-	-	-7.42	5.13
2018-09-05 15:23:01	0.01	0.12	0.96	-	-	-6.99	-0.52
2018-09-05 15:20:12	-0.02	0.14	0.96	-	-	-8.05	1.18
2018-09-05 15:14:37	-0.05	0.02	0.97	-	-	-1.21	2.86
2018-09-05 15:12:34	0.00	-0.43	0.88	-	-	25.99	0.08
2018-09-05 15:07:30	-0.09	0.01	0.97	-	-	-0.38	5.41
2018-09-05 14:58:17	-0.56	0.11	0.81	-	-	-6.12	34.49
2018-09-05 14:53:06	-0.73	0.15	0.67	-	-	-8.88	47.29
2018-09-05 14:42:45	0.06	0.05	0.97	-	-	-2.89	-3.58
2018-09-05 14:39:45	-0.02	0.01	0.97	-	-	-0.60	1.20

1 2 3 4

Data history for PyCom

- The transferred data is shown in a Grafana dashboard.

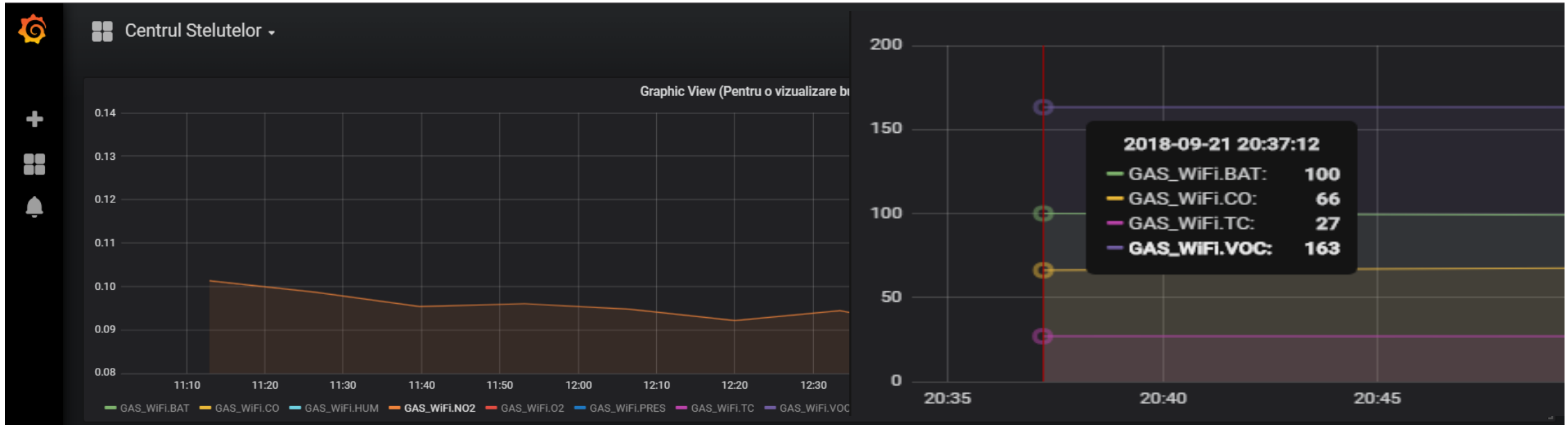
- The dashboard created for PyTrack in Grafana, the three axes of the accelerometer and the latitude and longitude. In this case, the pictures were taken with the device indoor, so the latitude and longitude can't be defined.

Libelium

- Libelium stations are utilized for analyzing:
  - the level of pollution - air pollution
  - agricultural parameters
  - noise
  - water
  - meteorology
  - water management systems
  - measurement of renewable energy potential plant disease management, etc.

Libelium Plug&Sense

Libelium

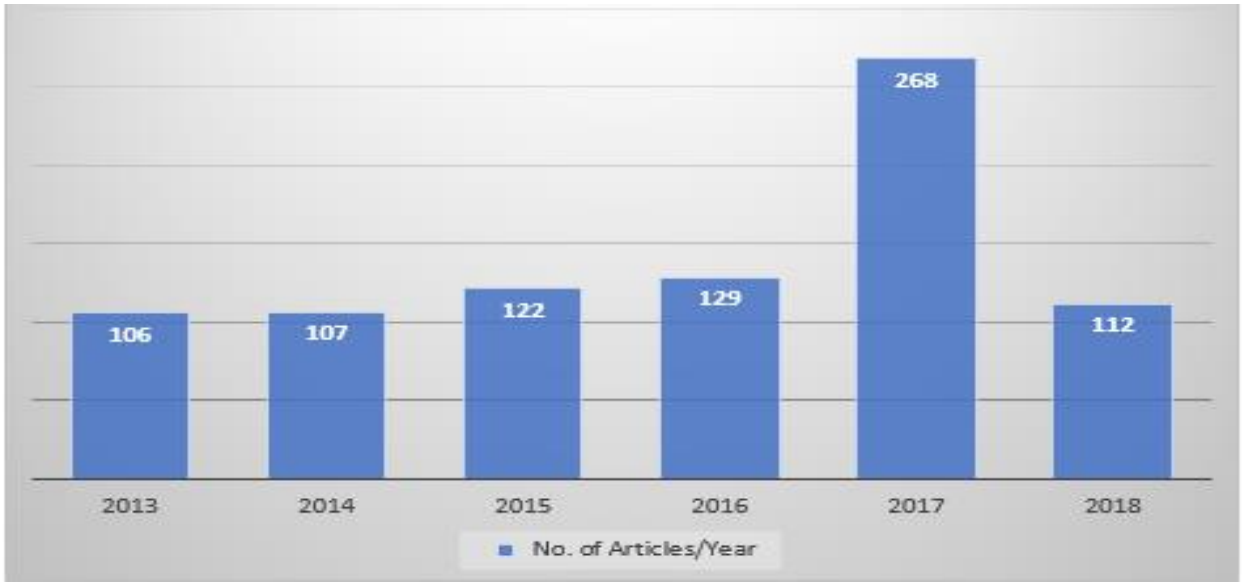


Graph based on data gathered from Libelium sensors

- Libelium stations can support different types of sensors. The data from these sensors are stored in the Libelium gateway, called Meshlium.
- we have connected three sensors (CO, Temperature, Volatile organic compound (VOC)) to a Libelium board

Big data for academic research

- Due to its notable potential in generating business value, Big Data has become the focus of academic and corporate investigation.
- The results of researching for all available journals and articles which are dealing with Big Data. In order to retrieve the articles, we searched by keywords using a Python script and Scopus library.



The distribution of articles related to big data

Conclusions

- We have presented how the users can manage with the huge amount of data gathered from different sensors which are connected to three different devices.
- Also, we have analyzed the corpus of different journals in order to identify the importance of Big Data for scientific research. This method is also useful for making a relevant database for scientific research and it also allows the user to identify the Regions of Interest (ROIs).

Future work

As future work:

- A relevant database using a Python script and Scholarly library which retrieves all articles available on Google Scholar will be created. The two Libraries will be compared.

Thank you!

Contact:
george@beia.ro