61-139 Poznan
ul. Jana Pawła II 10
phone: (+48 61) 858-20-01
fax: (+48 61) 852-59-54
office@man.poznan.pl
www.psnc.pl

Krzysztof Martyn

Anomaly detection in Data Center infrastructure

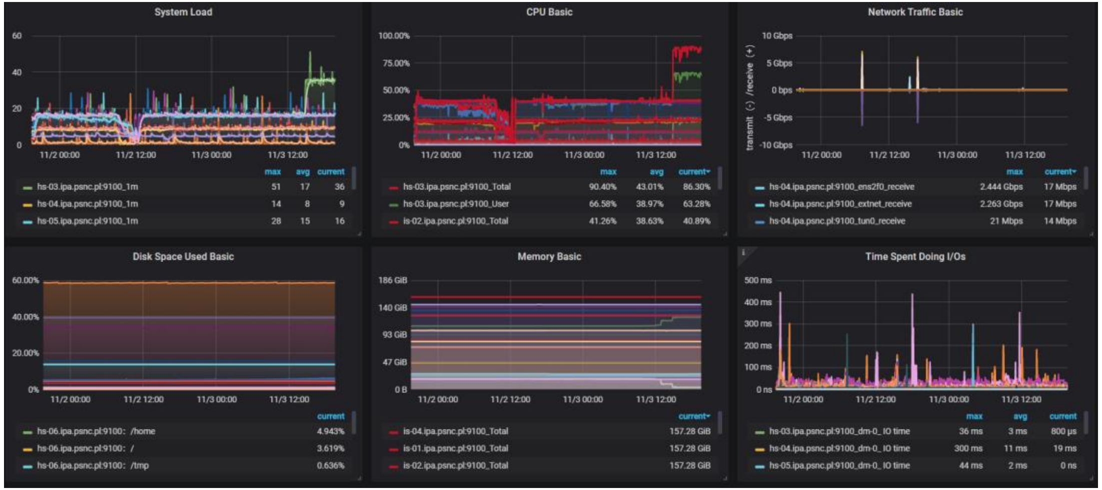
Data center monitoring

Basic approach:

- Controlling the current state and comparing it with the previous
- Look only of basic server metrics: **CPU, RAM, disk** and **network information**

Problems:

- The health of the system depends on all the components
- **Variety** of metrics
- **Heterogeneity** of monitored devices
- **Rapidly changing metrics**
- Changing characteristics of the server operation
- The need to **react quickly** to problems
- **Experience**



System Load

	max	avg	current
▬ hs-03.ipa.psnc.pl:9100_1m	51	17	36
▬ hs-04.ipa.psnc.pl:9100_1m	14	8	9
▬ hs-05.ipa.psnc.pl:9100_1m	28	15	16

CPU Basic

	max	avg	current▾
▬ hs-03.ipa.psnc.pl:9100_Total	90.40%	43.01%	86.30%
▬ hs-03.ipa.psnc.pl:9100_User	66.58%	38.97%	63.28%
▬ is-02.ipa.psnc.pl:9100_Total	41.26%	38.63%	40.89%

Network Traffic Basic

	max	current▾
▬ hs-04.ipa.psnc.pl:9100_ens2f0_receive	2.444 Gbps	17 Mbps
▬ hs-04.ipa.psnc.pl:9100_extnet_receive	2.263 Gbps	17 Mbps
▬ hs-04.ipa.psnc.pl:9100_tun0_receive	21 Mbps	14 Mbps

Disk Space Used Basic

	current
▬ hs-06.ipa.psnc.pl:9100: /home	4.943%
▬ hs-06.ipa.psnc.pl:9100: /	3.619%
▬ hs-06.ipa.psnc.pl:9100: /tmp	0.636%

Memory Basic

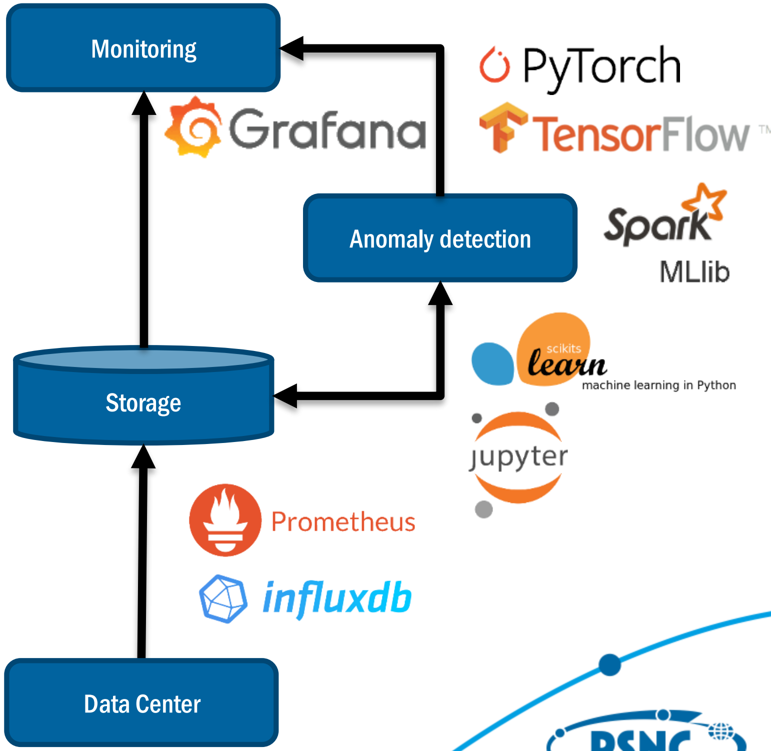
	current▾
▬ is-04.ipa.psnc.pl:9100_Total	157.28 GiB
▬ is-01.ipa.psnc.pl:9100_Total	157.28 GiB
▬ is-02.ipa.psnc.pl:9100_Total	157.28 GiB

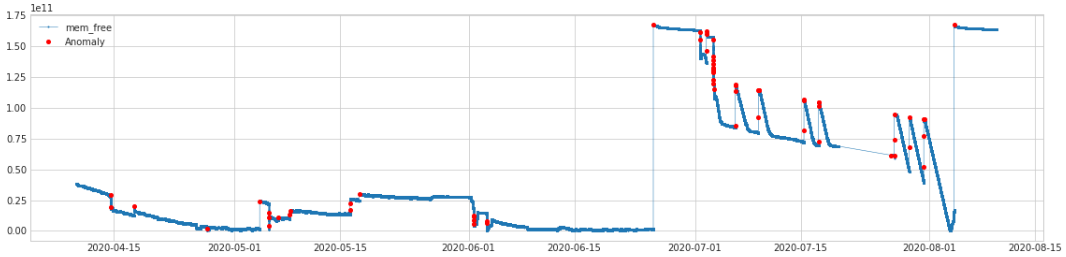
Time Spent Doing I/Os

	max	avg	current
▬ hs-03.ipa.psnc.pl:9100_dm-0_ IO time	36 ms	3 ms	800 µs
▬ hs-04.ipa.psnc.pl:9100_dm-0_ IO time	300 ms	11 ms	19 ms
▬ hs-05.ipa.psnc.pl:9100_dm-0_ IO time	44 ms	2 ms	0 ns

Anomaly detection

1. Data collected from each server
2. Data aggregation from multiple devices
3. Online anomaly detection using multiple independent machine learning, deep learning and statistical methods
4. Automatic alerts when anomalies are detected
5. Displaying the current and historical data along with detected anomalies

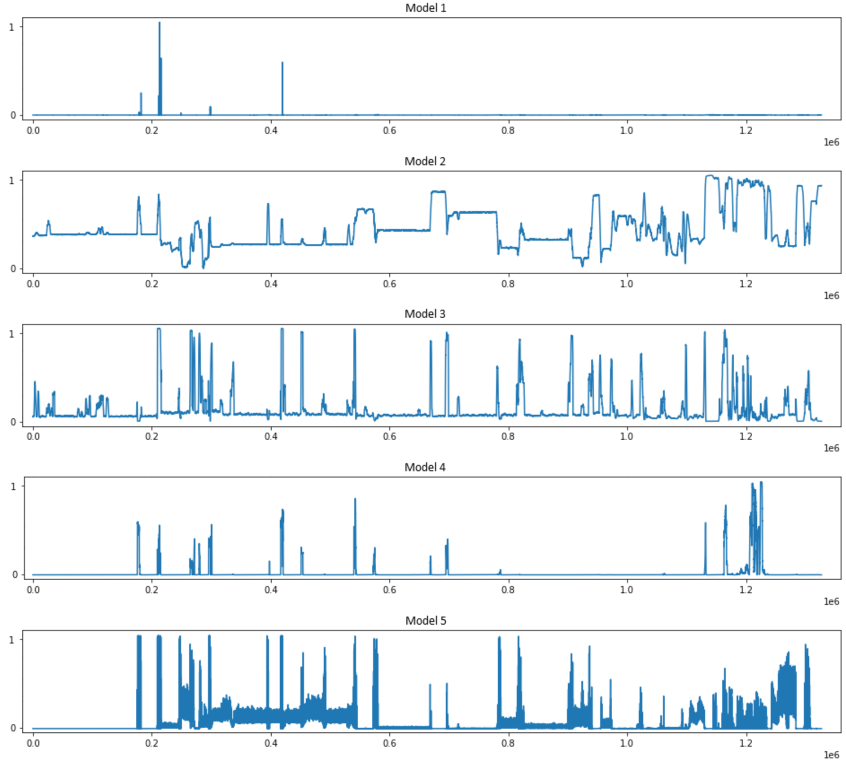




Anomaly detection with machine learning

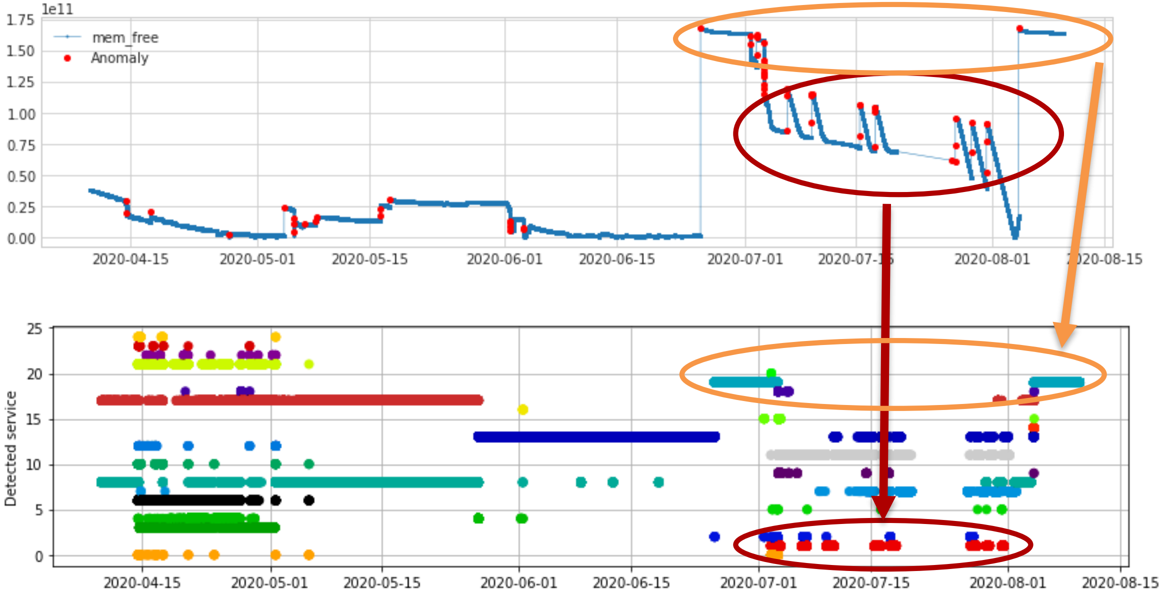
- Discovering the usual characteristics of devices
- Profiling the behavior of each server separately to discover the individual behavior of each server
- Modeled the state of the entire system and comparing it with historical data

- Use of various anomaly detection techniques:
  - **Predictive models** - an anomaly if the data does not agree with the prediction
  - **Statistical models** - an anomaly if the dynamics of changes is inconsistent with the previous one
  - **State models** - an anomaly if the data corresponds to the state in which server should not currently appear
  - **Threshold** baseed methods - an anomaly if any metric exceeds the threshold



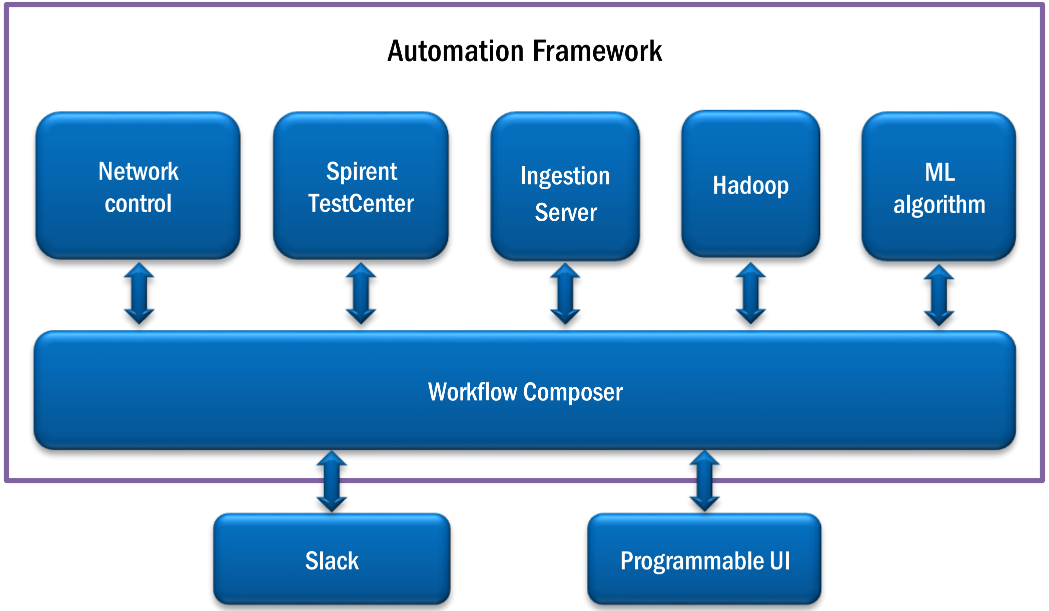
Anomaly detection with machine learning

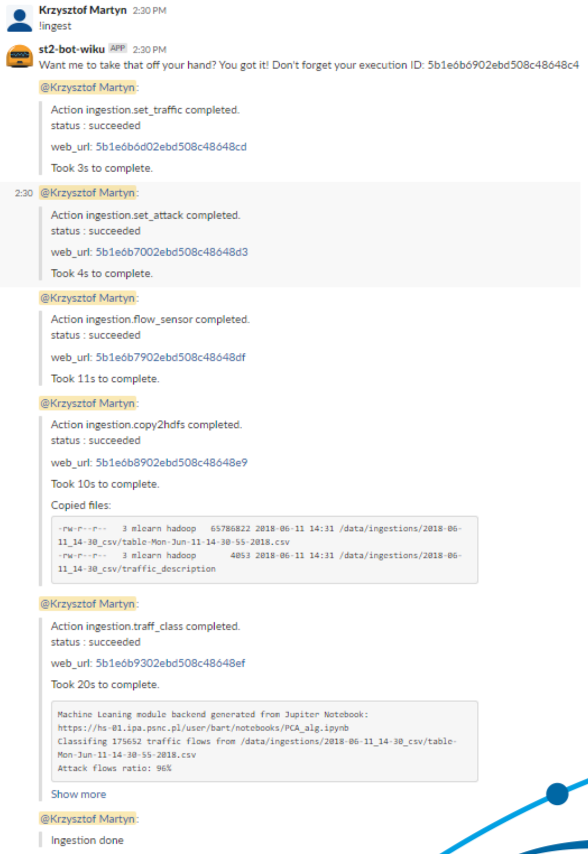
- Profiling the server behavior
- Detection of running services on the server by analyzing the similarity of behavior



Platform for the network ML

- Real traffic from the POZMAN network
- Ability to inject an artificial attack generated by Spirent TestCeneter
- Jupyter Notebooks for researchers to develop ML prototypes
- A Slack account integrated with Automation Tools
- Usage example: DDoS attack detection





Krzysztof Martyn 2:30 PM
!ingest

st2-bot-wiku APP 2:30 PM
Want me to take that off your hand? You got it! Don't forget your execution ID: 5b1e6b6902ebd508c48648c4

@Krzysztof Martyn:

Action ingestion.set_traffic completed.
status : succeeded

web_url: 5b1e6b6d02ebd508c48648cd

Took 3s to complete.

2:30 @Krzysztof Martyn:

Action ingestion.set_attack completed.
status : succeeded

web_url: 5b1e6b7002ebd508c48648d3

Took 4s to complete.

@Krzysztof Martyn:

Action ingestion.flow_sensor completed.
status : succeeded

web_url: 5b1e6b7902ebd508c48648df

Took 11s to complete.

@Krzysztof Martyn:

Action ingestion.copy2hdfs completed.
status : succeeded

web_url: 5b1e6b8902ebd508c48648e9

Took 10s to complete.

Copied files:

```
-rw-r--r--   3 mlearn hadoop   65786822 2018-06-11 14:31 /data/ingestions/2018-06-
11_14-30_csv/table-Mon-Jun-11-14-30-55-2018.csv
-rw-r--r--   3 mlearn hadoop       4053 2018-06-11 14:31 /data/ingestions/2018-06-
11_14-30_csv/traffic_description
```

@Krzysztof Martyn:

Action ingestion.traff_class completed.
status : succeeded

web_url: 5b1e6b9302ebd508c48648ef

Took 20s to complete.

```
Machine Leaning module backend generated from Jupiter Notebook:
https://hs-01.ipa.psnc.pl/user/bart/notebooks/PCA_alg.ipynb
Classifing 175652 traffic flows from /data/ingestions/2018-06-11_14-30_csv/table-
Mon-Jun-11-14-30-55-2018.csv
Attack flows ratio: 96%
```

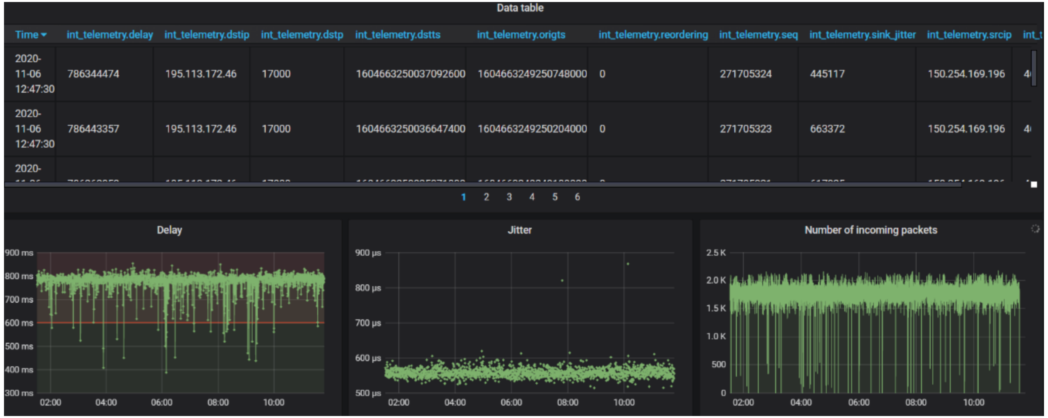
Show more

@Krzysztof Martyn:

Ingestion done

Future

- Integration with **NMaaS**
- Analize **In-band Network Telemetry**



Data table

Time	int_telemetry.delay	int_telemetry.dstip	int_telemetry.dstp	int_telemetry.dstts	int_telemetry.origts	int_telemetry.reordering	int_telemetry.seq	int_telemetry.sink_jitter	int_telemetry.srcip	int_t
2020-11-06 12:47:30	786344474	195.113.172.46	17000	1604663250037092600	1604663249250748000	0	271705324	445117	150.254.169.196	4
2020-11-06 12:47:30	786443357	195.113.172.46	17000	1604663250036647400	1604663249250204000	0	271705323	663372	150.254.169.196	4
2020-										

Poznan Supercomputing and Networking Center

61-139 Poznan
ul. Jana Pawła II 10
phone: (+48 61) 858-20-01
fax: (+48 61) 852-59-54
office@man.poznan.pl
www.psnc.pl