```
Capybara.javascript_driver =
Category.delete_all;
Shoulda::Matchers.configure
  config.integrate do
    with.test_framework
    with.library :rails
  end
end

# Add additional requires

# Requires supporting
```

Anomaly Detection on Sensor Data of a Deep Learning Development Server Cluster

Project Goal

Project goal was to detect anomalies of sensor data of a deep learning development server cluster. The cluster consists of high scaled GPU servers and is utilized to process high volumes of data to train machine learning models and neural networks. These kind of processes run normally for several hours. The servers are accessed by many different developers working remotely. This makes them a bottleneck resource. To gain more information about the general occupancy of the cluster and to get notifications when big training processes are kicked off, a monitoring and anomaly detection solution was needed. Like this automated alerts for major deviation from normal behaviors should be reported to the person in charge via e-mail.

Provided Data

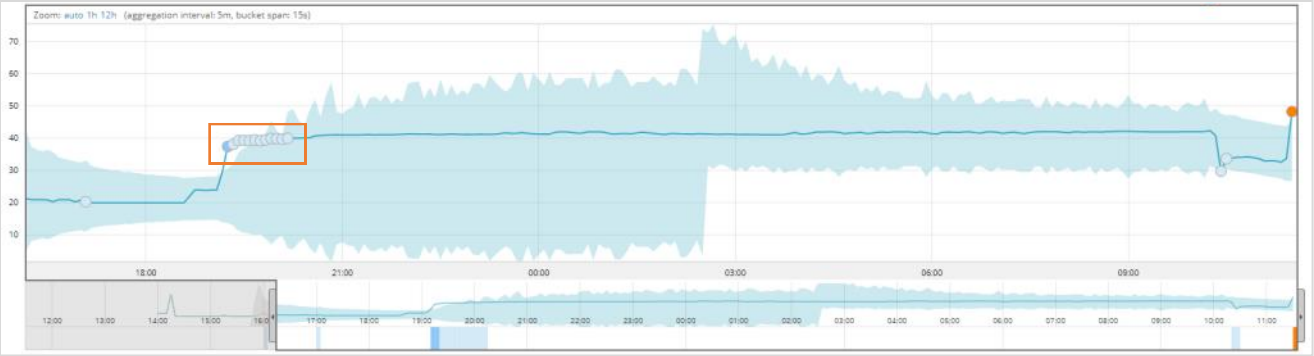
For each server 58 features are recorded every second and stored in a CSV file separated by server. The CSV files are split every 24 hours to avoid large files, as more than 65,000 lines are produced each day.

The sensors register several data types for the different hardware components. Differing from capacity, temperature, clock speeds and power consumption they cover all important aspects to evaluate the current workloads and occupancy rates.

Challenges

The data is recorded and stored on the development servers, while the analytics tool runs on another dedicated server. This data must be made available for the analysis server while keeping the transfer delays at a minimum.

As the deep learning servers will mostly have either moderate or high capacities (depending on processing a model or running idle), the alerts can not just be configured for certain threshold values. The machine learning model needs to recognize and incorporate trends and notice unexpected trend changes or peaks - that means: the model needs to adjust itself over time.



Example of detected anomalies and self adjusting trend boundaries

Applied Methods

The chosen analytics tool is the highly scalable Elastic Stack. With the Logstash module a real-time pipeline was build up, to directly store the most current data from the log files in an Elasticsearch cluster. Kibana was set up as the interface for the machine learning module, which is part of the X-Pack extension.

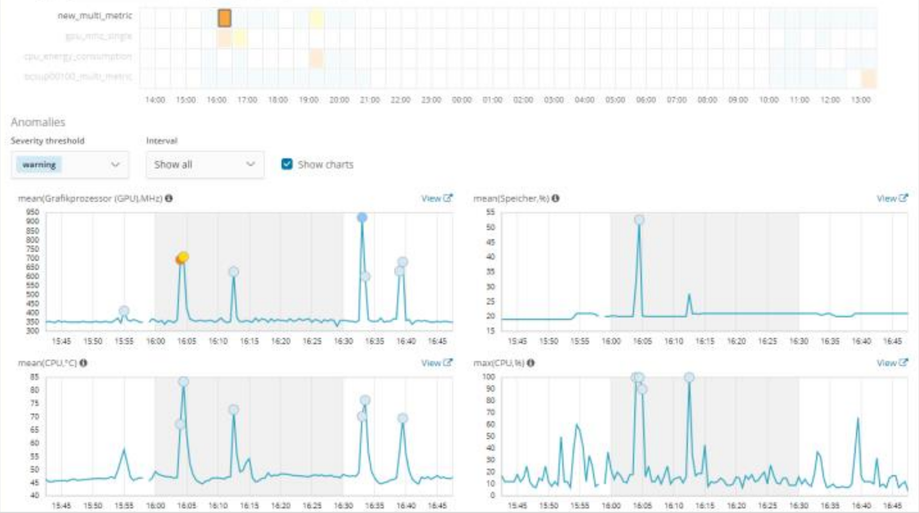
To make the log data available, direct network access for the log file directories was configured. Like this the logs can be directly accessed by Logstash, which is permanently checking for new information and uploads it directly into Elasticsearch. The most current sensor data is ready to be analyzed with less than a second of delay.

The Machine Learning Module was configured to monitor several critical sensors and inference trends in the data. It analyzes temperature, clock speeds and capacities for CPU, GPU, RAM and storage components.

To notify the responsible administrators about exceptional events on their server cluster, e-mail alerts for different severity thresholds were configured.

Project Outcome

A real-time monitoring tool with a self learning anomaly detection model was implemented. It is configured to automatically send e-mails, for certain severity thresholds of several hardware components. Like this the utilization of the deep learning server cluster can be managed more effectively. The monitoring also provides a better overview of the capacities and potential hardware limitations the developers might run into.



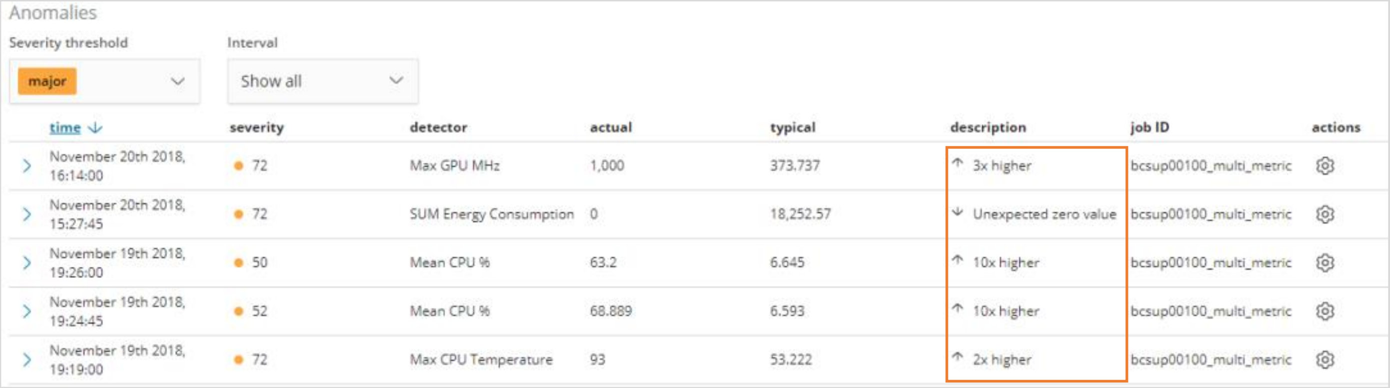
Anomalies

Severity threshold: warning

Interval: Show all

☑ Show charts

The Anomaly Explorer gives an overview of all monitored metrics



Anomalies

Severity threshold: major

Interval: Show all

	time ↓	severity	detector	actual	typical	description	job ID	actions
›	November 20th 2018, 16:14:00	● 72	Max GPU MHz	1,000	373.737	↑ 3x higher	bcsup00100_multi_metric	⚙
›	November 20th 2018, 15:27:45	● 72	SUM Energy Consumption	0	18,252.57	↓ Unexpected zero value	bcsup00100_multi_metric	⚙
›	November 19th 2018, 19:26:00	● 50	Mean CPU %	63.2	6.645	↑ 10x higher	bcsup00100_multi_metric	⚙
›	November 19th 2018, 19:24:45	● 52	Mean CPU %	68.889	6.593	↑ 10x higher	bcsup00100_multi_metric	⚙
›	November 19th 2018, 19:19:00	● 72	Max CPU Temperature	93	53.222	↑ 2x higher	bcsup00100_multi_metric	⚙

List of detected anomalies for one of the servers

Further Applications

The machine learning module has an integrated calendar tool to specify events, which cause peaks in the data but should not be classified as anomalies. In the near future developers will be able to book time slots for using the development servers in this calendar. Alerts will only be sent out of those time frames. This lowers the number of alerts for planned activities on the server and helps to administrate the planning process for parallel running projects.

In addition the alerts can be configured more precisely with certain rules. More of those rules will be incorporated in the testing phase of the system to increase the precision of the alerts and lower the number of false alerts.