System for visualizing resource usage statistics of a computing cluster

Maksim Lebedev, Danil Dereza, FEFU, Vladivostok
Dmitry Belyakov, Maksim Zuev, MLIT JINR, Dubna
AYSS, JINR
2024

HybriLIT Platform & Govorun Supercomputer

HybriLIT

HybriLIT is a hybrid computing platform providing infrastructure for high-performance computing and is part of the Multifunctional Information and Computing Complex (MICC) of the Laboratory of Information Technologies.

Govorun

Govorun Supercomputer is located at JINR and is designed to solve tasks in scientific research.

Relevance of the System

The availability of a specialized statistics system will make it possible to conduct detailed performance analysis and use historical data for optimization of computing cluster, which reduces resource costs and allows to anticipate possible problems in advance.

The project should close the existing needs for visualization of statistics and improve the quality of analysis of computing processes, which is extremely important in conditions of highly loaded computing centers.

Salsa

 Grafana

Goal and Objectives of the System

■ **GOAL**

Creation an effective system for analyzing and visualizing statistical data collected by the HybriLIT computing node monitoring system using the Yandex DataLens BI platform.

■ **OBJECTIVES**

- Deployment of the Yandex DataLens BI platform for access from the JINR local network
- Organization of user authorization using FreeIPA and Authelia
- Aggregation of data obtained by the monitoring system
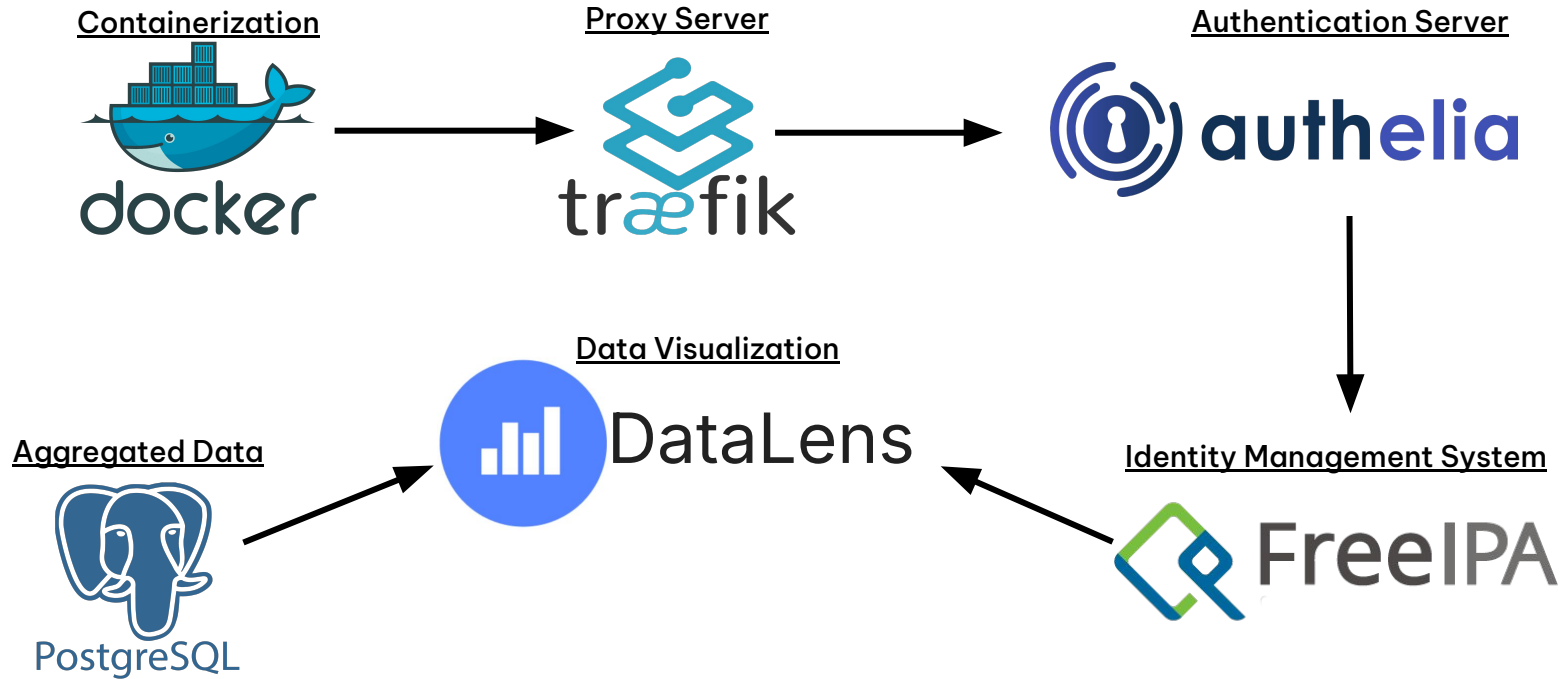- Building informative dashboards for statistical research

Yandex DataLens

Yandex DataLens is a Business Intelligence tool from Yandex that allows to visualize data and build reports in real time.

The main functions of DataLens are aimed at analyzing data from various sources, creating visual visualizations and simplifying the decision-making process.
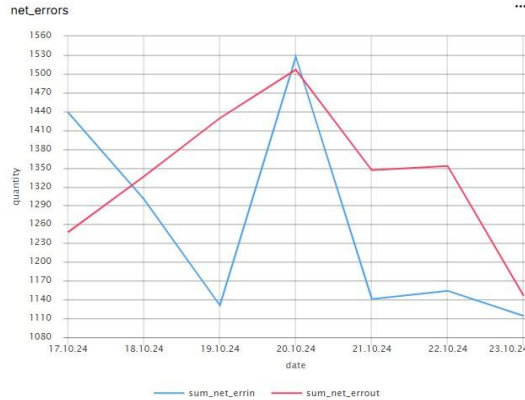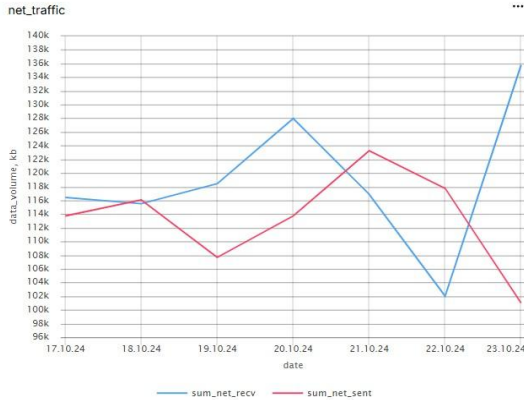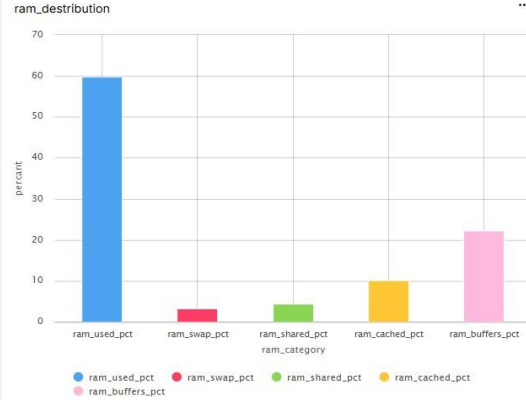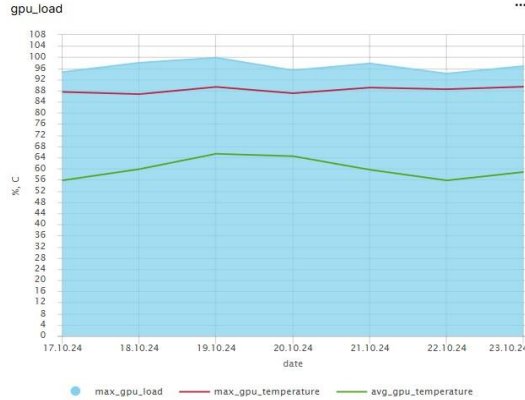
System Architecture

Containerization
docker

Proxy Server
træfik

Authentication Server
authelia

Data Visualization
DataLens

Aggregated Data
PostgreSQL

Identity Management System
FreeIPA

Interaction with the Resource Monitoring System Database

ALGORITHM FOR COLLECTING DATA

Monitoring System → Timescale → PostgreSQL

For each node of the computing cluster, data with different periods are regularly taken from the TimescaleDB: for the last week, month and year. This data is then written to PostgreSQL at 5-minute intervals. After that, they are unloaded from PostgreSQL and sent to DataLens for visualization.

Resource Usage Visualization System: Dashboards

Results / Current Advantages
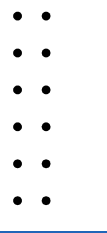
ACHIEVED RESULTS

- Specialized system for collecting and analyzing of statistics for HybriLIT Platform was developed
- It was deployed on the Heterogeneous platform HybriLIT

ADVANTAGES

- Detailed historical analysis
- Flexibility and adaptability
- Decision-making process
- Convenient data visualization

Further System Improvement Plans

- Adding authentication

- Importing real data

- Adding a user manual

- Adding of charts at the request of members of the HybriLIT group

Thank you for attention!

Special thanks
HybriLIT heterogeneous group
University Center, JINR