

PMP Data Analysis

The GN5-2 Network Development Work Package – WP6

Krešimir Šodan & Ljubomir Hrboka

GN5-2 Work Package 6 Task 3

Monitoring and Management

5th European perfSONAR User Workshop

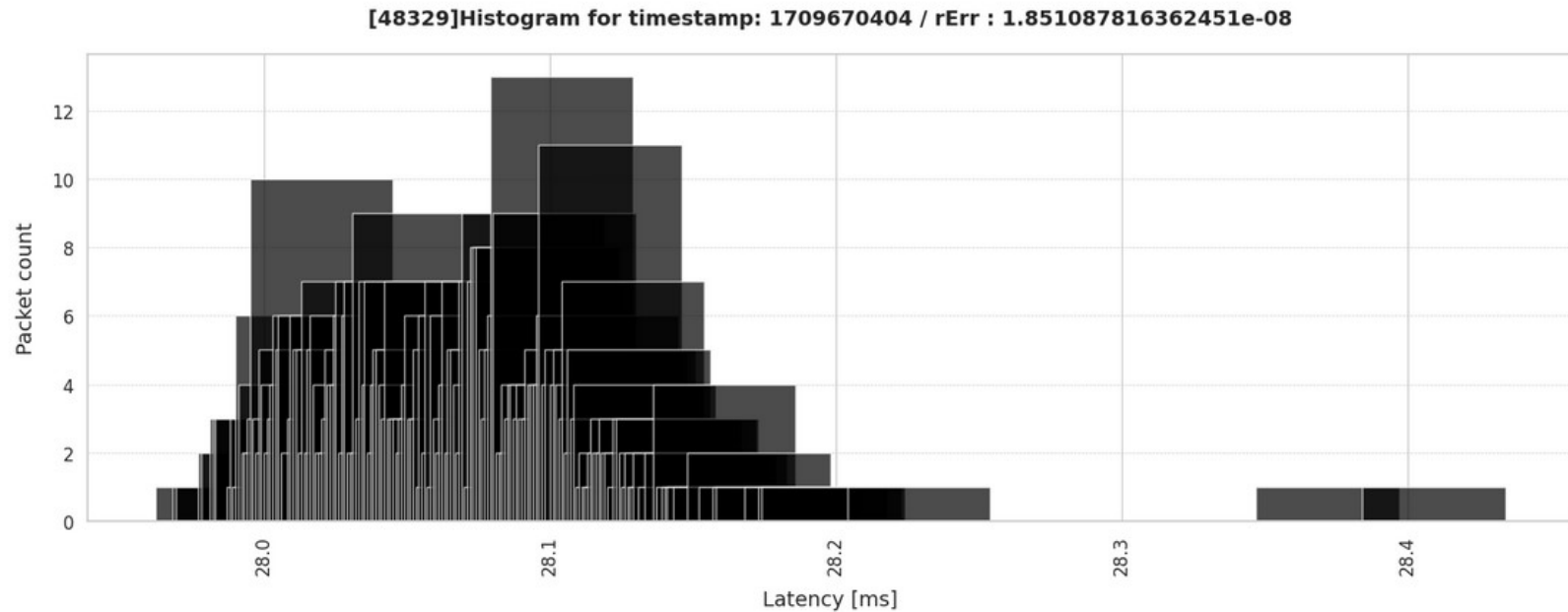
6th - 7th May 2026 Paris, France

Public (PU)

Introduction

- Monitoring systems rely on static thresholds
 - subtle anomalies
 - gradual performance degradation
- Machine learning
 - alternative approach
 - correlation with other monitoring tools
- PerfSONAR OWD measurement
 - Frequency – 60 seconds
 - 1440 measurements a day

Data



Data preprocessing

- Standardization of input format
 - Consistent
 - Comparable
 - Structured
- Preprocessing pipeline:
 - Histogram resampling (256 bins)
 - Normalization – range [0,1]
 - Noise Handling – removing measurement error values

[illegible]

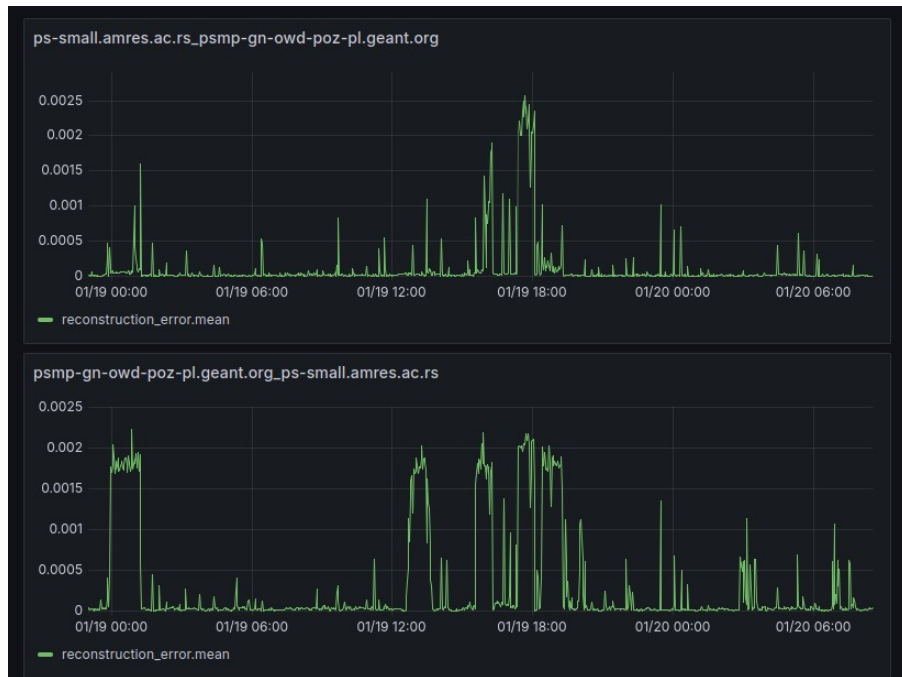
Machine Learning Model

- Unsupervised machine learning – no labelled data
- Trained on data representing normal network behaviour (~3 months)
- Autoencoder model:
 - Encoder – compress input data into a lower-dimensional representation
 - Decoder – reconstructs the original input
 - Reconstruction error as threshold

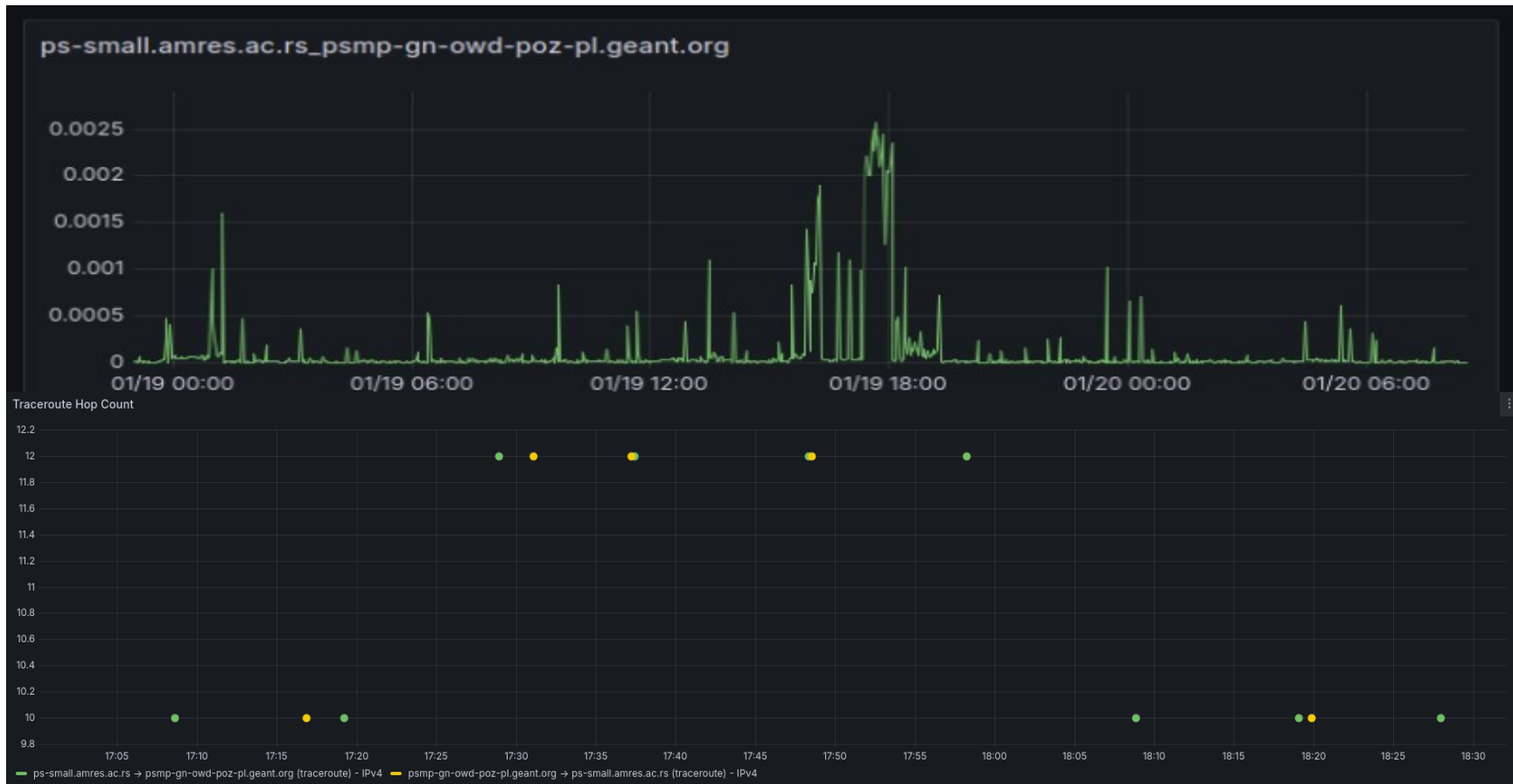
Architecture

- **Model training:**
 - Retrieve ~3 months of data per node pair & direction
(A->B , B->A)
 - Preprocessing
 - Train model on ~ 140,000 datapoints, save for inference
- **Inference:**
 - Retrieve latest OWD datapoint from Opensearch database
 - Preprocessing
 - Load designated model based on source and destination
 - Calculate reconstruction rate
 - Logging and plotting (Grafana dashboard)
 - Near real-time detection (every minute)

Analysis and results



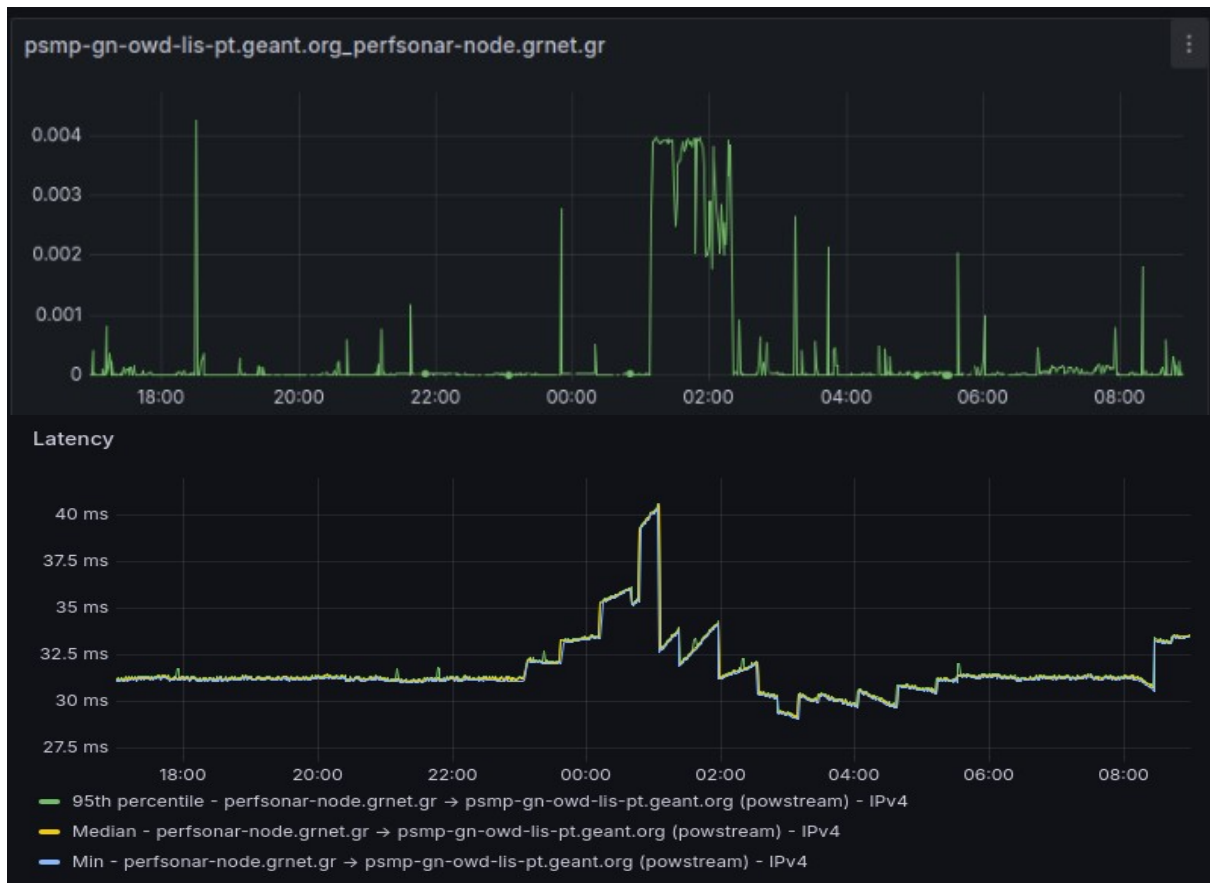
Analysis and results



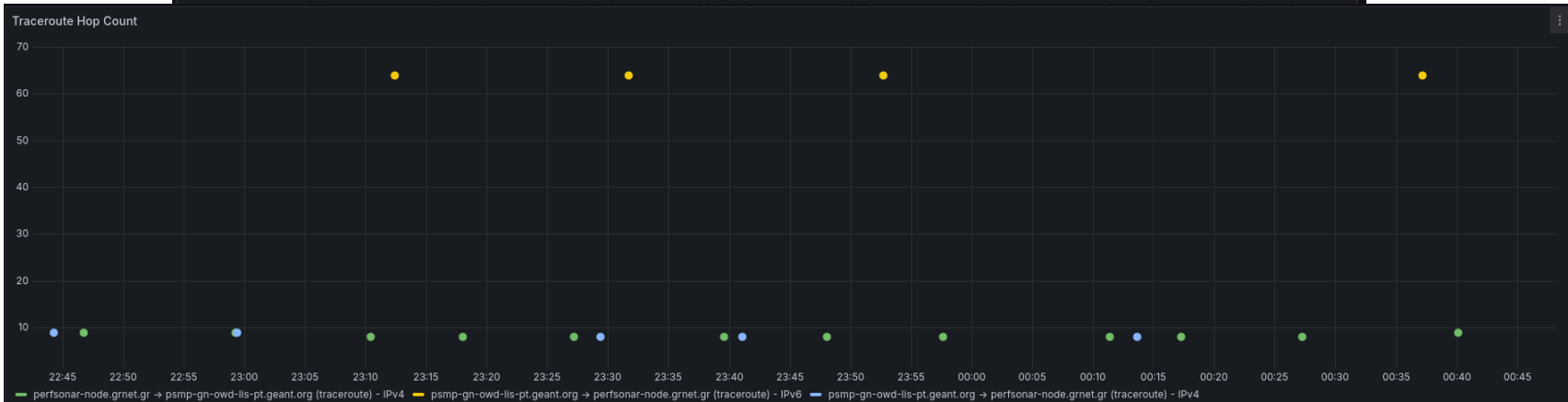
Analysis and results



Analysis and results



Analysis and results

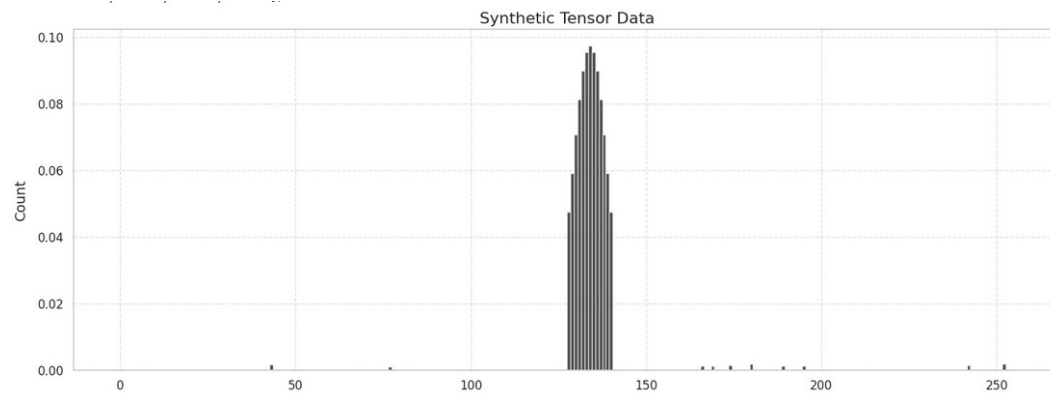
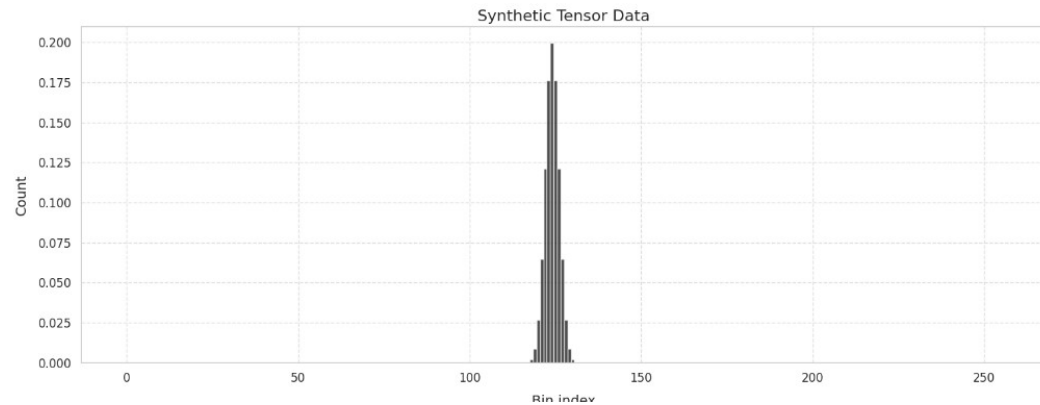


Directional differences

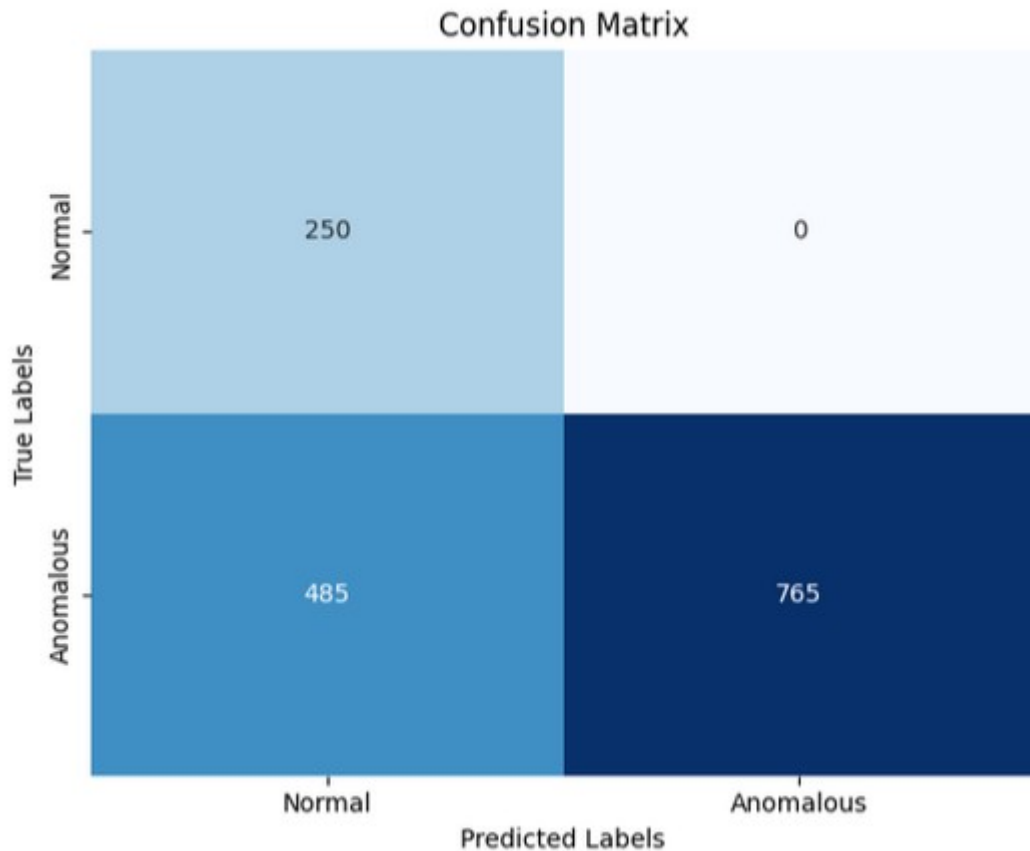
Anomaly patterns present in both directions

- Difference in magnitude and temporal properties
- Reflect same underlying event
- Forward and reverse path difference – asymmetry
- Correlation between directions – additional confidence

Synthetic data validation



Synthetic data validation



Conclusion

- Model sensitive to changes in the structure of latency distribution
- Detects less apparent network behaviour
- Normal behaviour vs fixed thresholds
- High precision (synthetic data validation)
- Complement to classical network monitoring



Thank You!

www.geant.org



Co-funded by
the European Union

Kresimir.sodan@carnet.hr