Data plane: Stop sampling and start scoring!

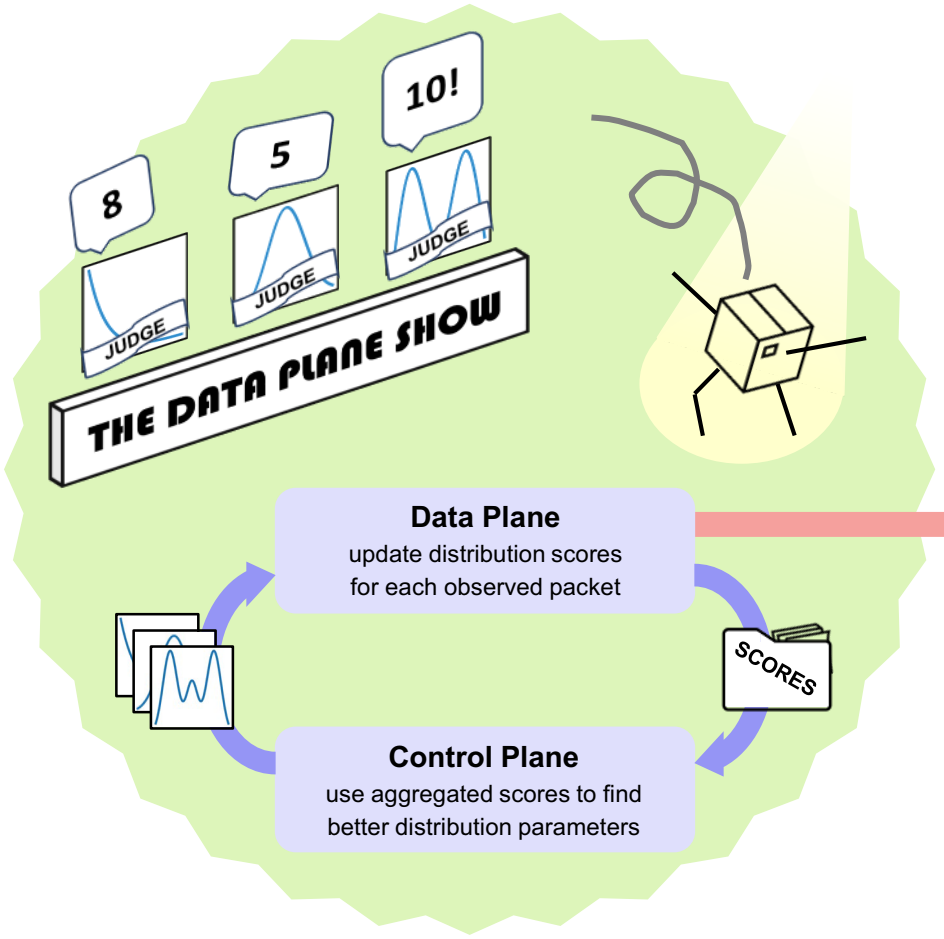
How to learn distributions that fit all traffic

Data plane: Stop sampling and start scoring!

How to learn distributions that fit *all* traffic.

The data plane sees all traffic, the control plane can run complex algorithms. **We need both!**

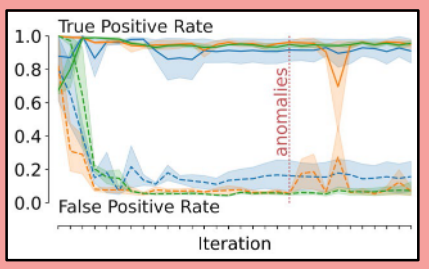
A feedback loop between planes: the control plane fits distributions using data-plane scores.



Don't let anomalies be the new normal!

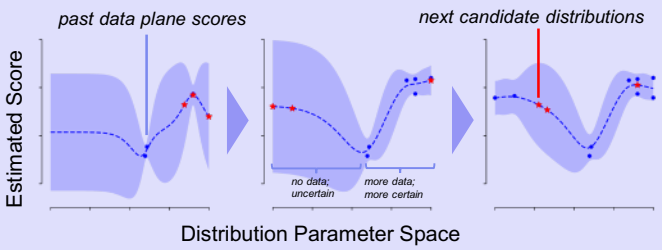
Anomaly detection is an important use-case: learned distributions are "normal" traffic. **Anomalies are identified in the data-plane from scores outside a confidence interval.**

Detection deteriorates if we wrongly learn that anomalies are normal (**orange**). We prevent this by filtering out anomalous scores (**green**).
*We show sampling as a baseline (**blue**) and mean/std from three random experiments.*



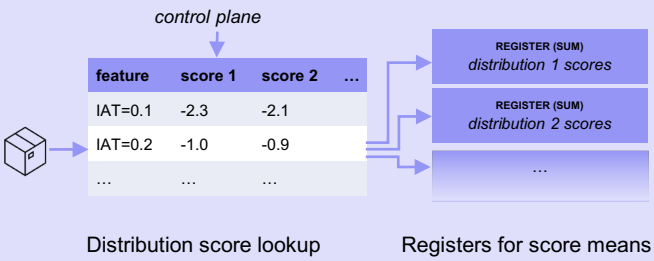
Control plane learning…

Bayesian optimization uses the (noisy) data plane scores to learn which distributions best fit the traffic.



…enabled by the data plane!

Proper scoring rules tell us which candidate fits the real traffic best and can be collected in the data plane.



WHAT NEXT?

We assume that anomalies are sharp changes while benign shifts are gradual.
Does this assumption hold in practice?

The number of parameters we can learn with Bayesian Optimization is limited.
Can we combine our approach with sampled data?


Alexander Dietmüller


Dr. Georgia Fragkouli


Prof. Dr. Laurent Vanbever

ETH zürich