Understanding Performance and its Anomalies in a VCL Cloud Using Rare Event Analysis

MŰEGYETEM 1782
Budapest University of Technology and Economics
Department of Measurement and Information Systems

Overview

Objective: operation support

- Highly dynamic system: diversity of workload – burstiness
- High availability → rare fault events

Exploratory Data Analysis – easy-to-use for operators

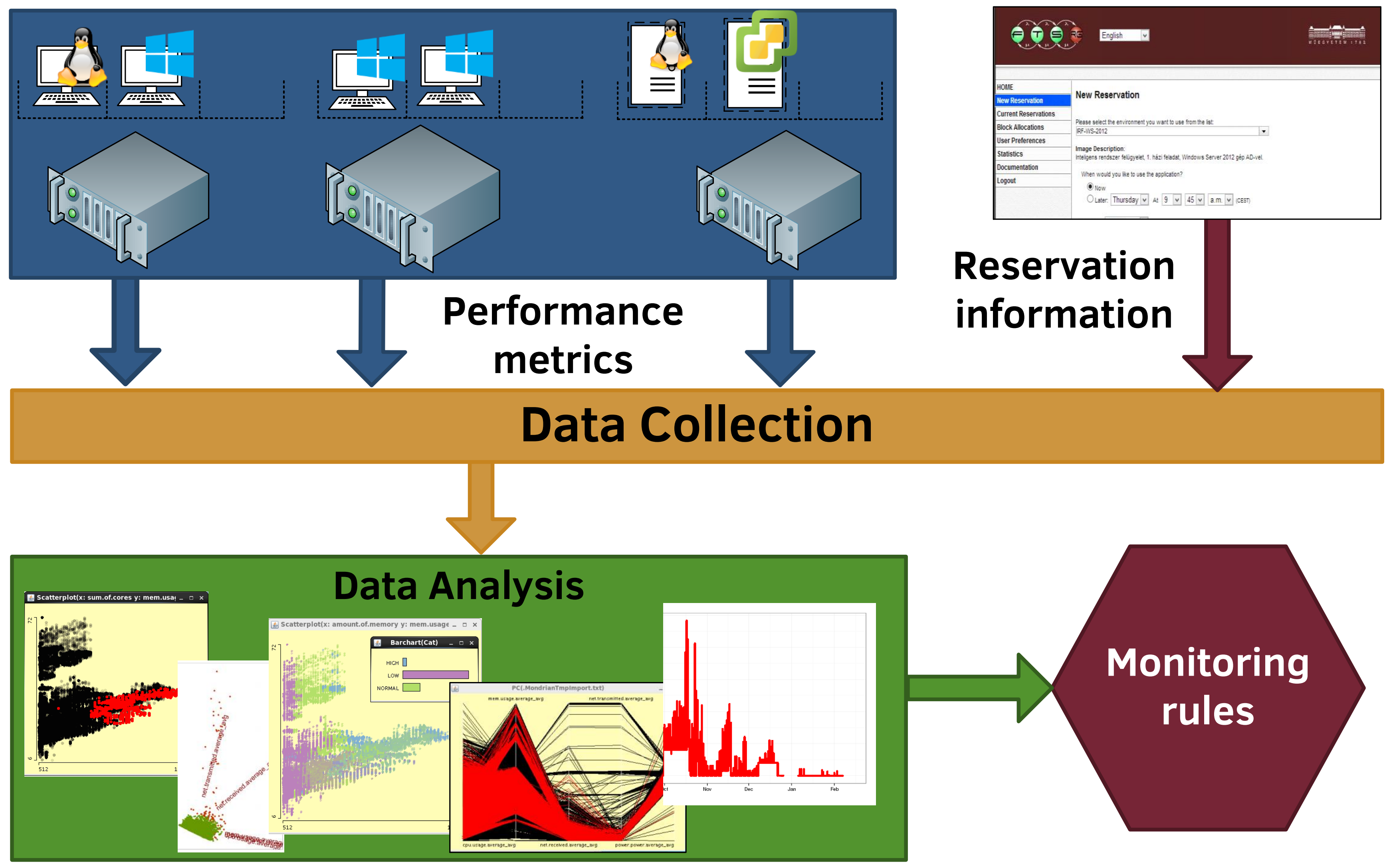
- Monitoring rules
- Capacity planning
- Risk modelling

Motivation

Observations:
VCL reservation + platform logs
40 metrics (20 s sampling period)
≈57 billion data points / year

- What to supervise?
- Early detection of deviances
- Long-term trends in resource utilization

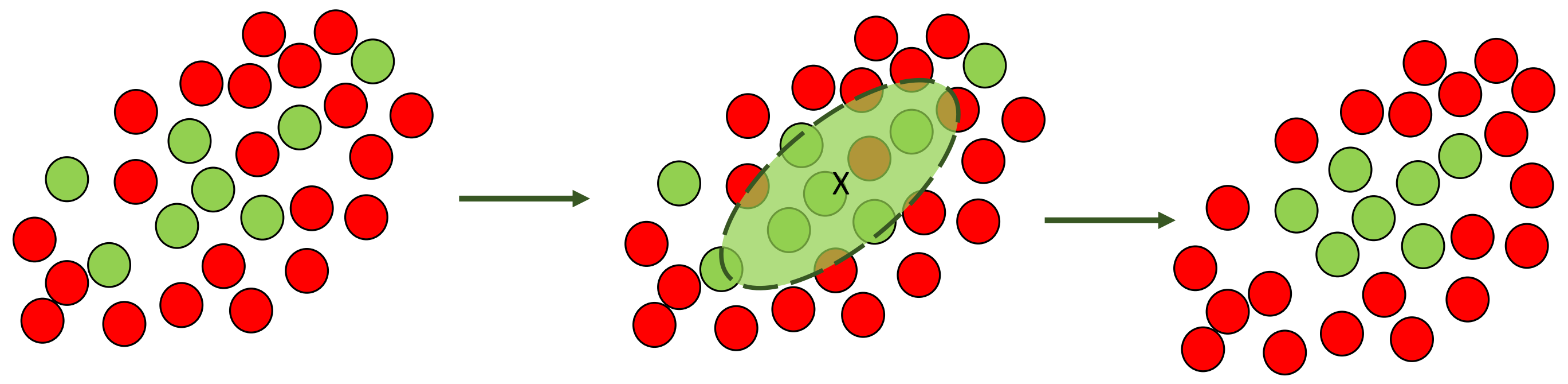
Hypervisor monitoring in VCL



Rare event identification

No a priori knowledge:

- System behaviour? Mixed workload, large and asynchronous user population
- Statistics of the phenomena?
- Uniform handling of different physical quantities observed? (Performance, number of VMs running)
- Learning/checking
- Algorithm: non-parametric, robust (BACON)



Characteristics:

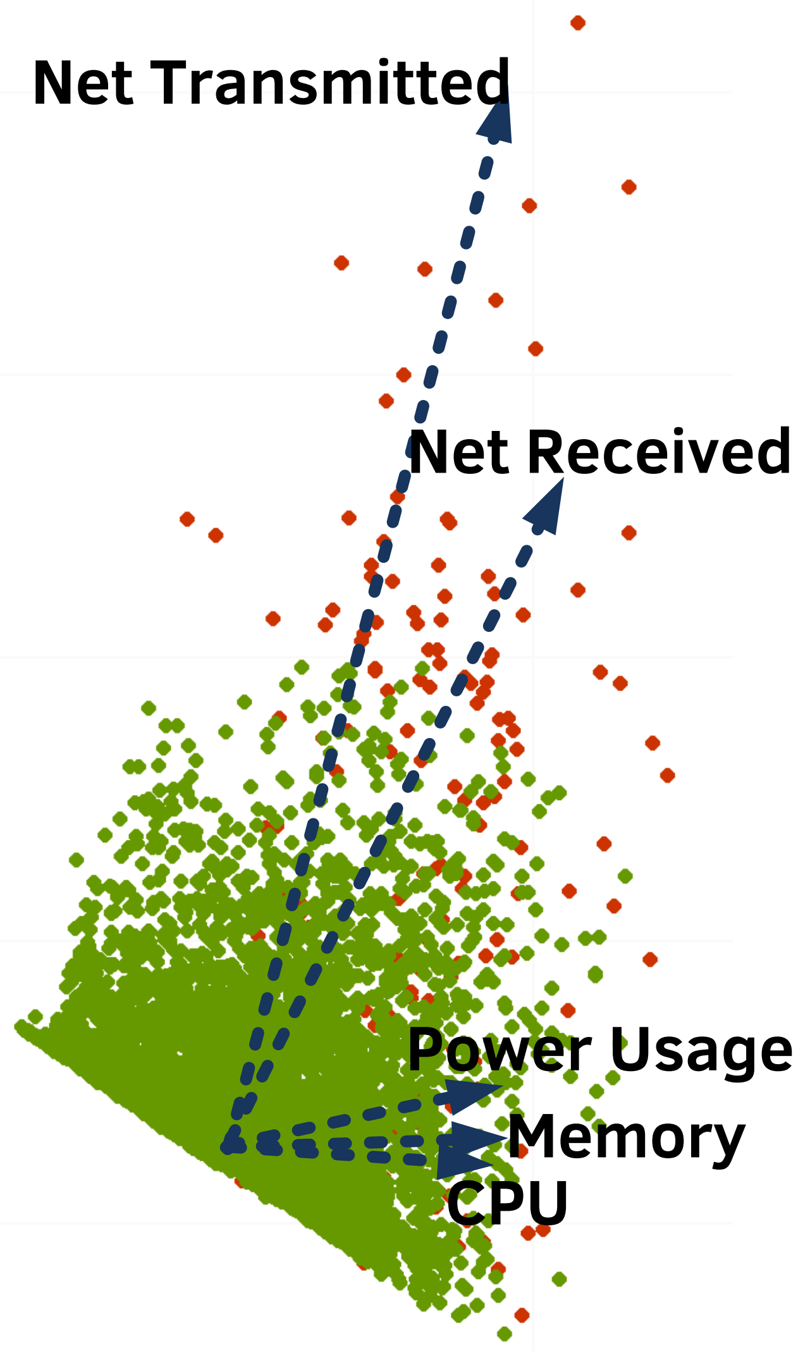
- Intelligent learning: semi-supervised/unsupervised mode
- Gradual refinement: iterative
- Mixed dimensions: Mahalanobis distance

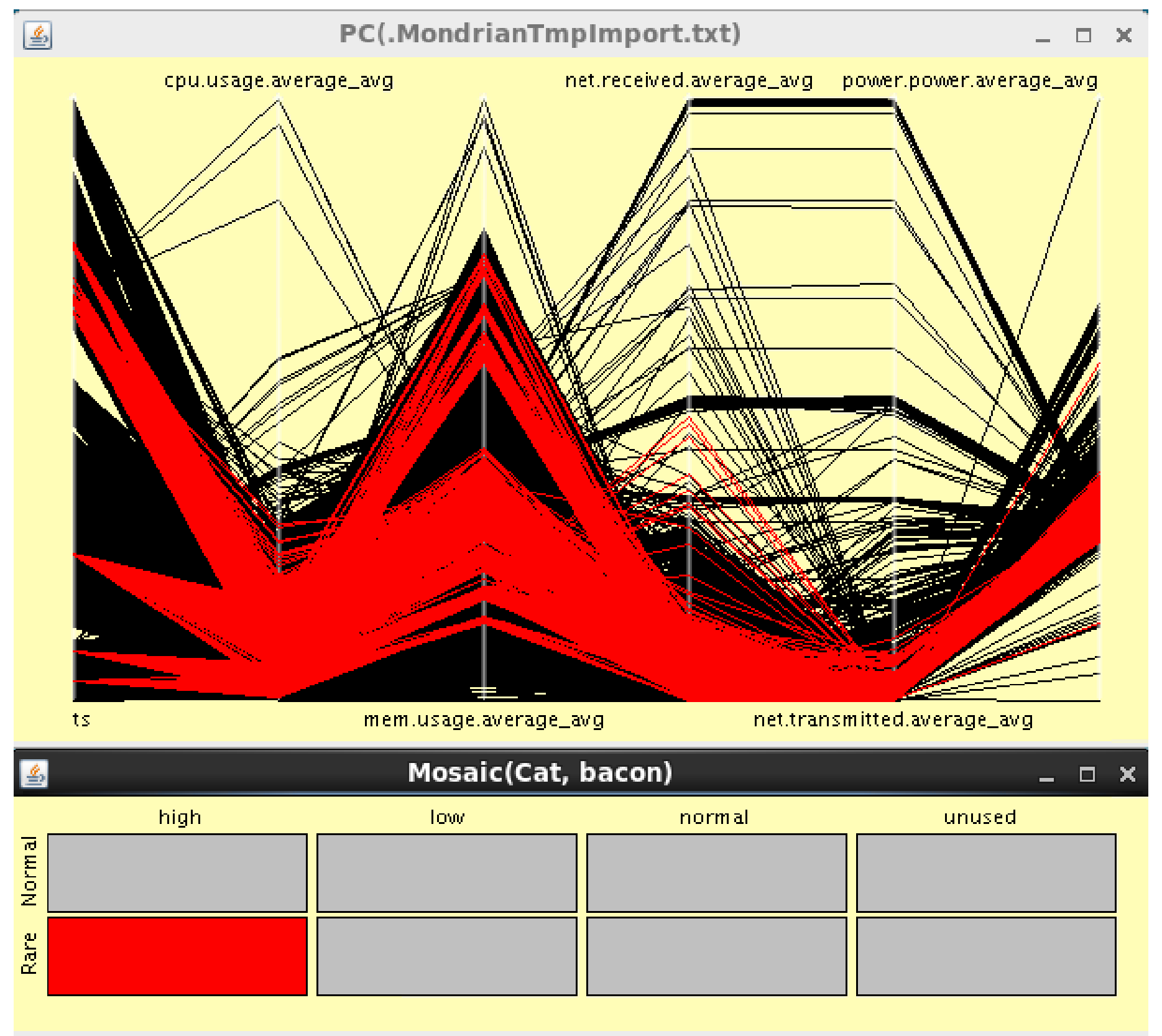
Performance in different operational domains

High workload category

"Faulty" domain

- Temporal burstiness → good predictability
- Rare events: dense clusters in all dimensions



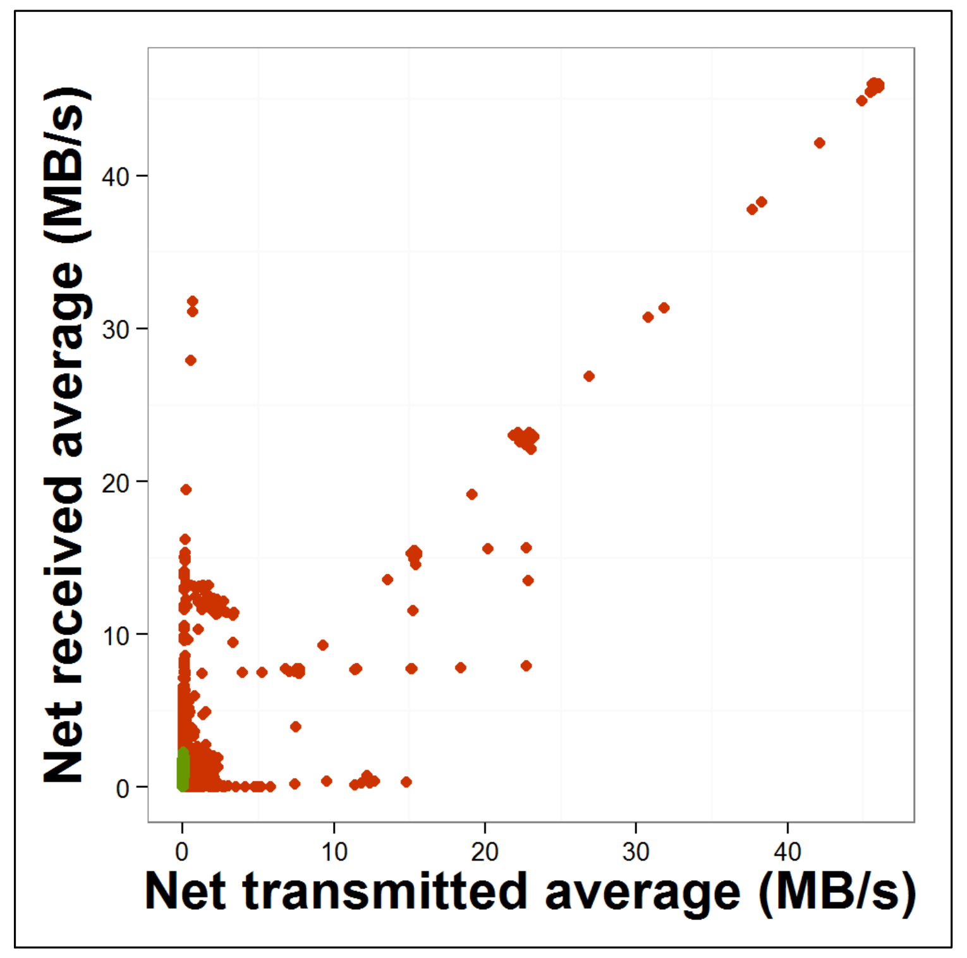


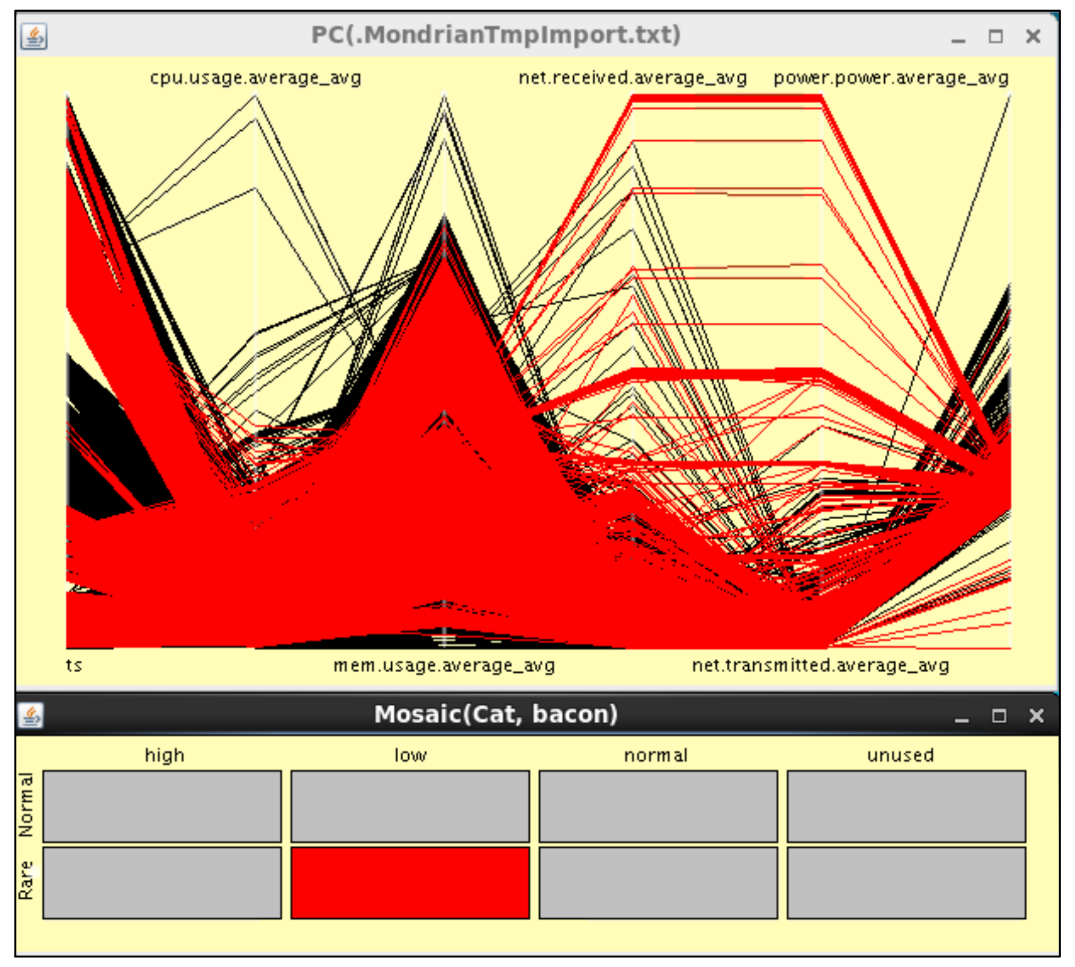
Normal domain

- Smoothing: > 20 VMs running simultaneously → strong smoothing
- Rare events: really extreme
- Data points: large difference in transmission – reception → suspicious
- Memory and CPU metrics: same trend

Low workload category

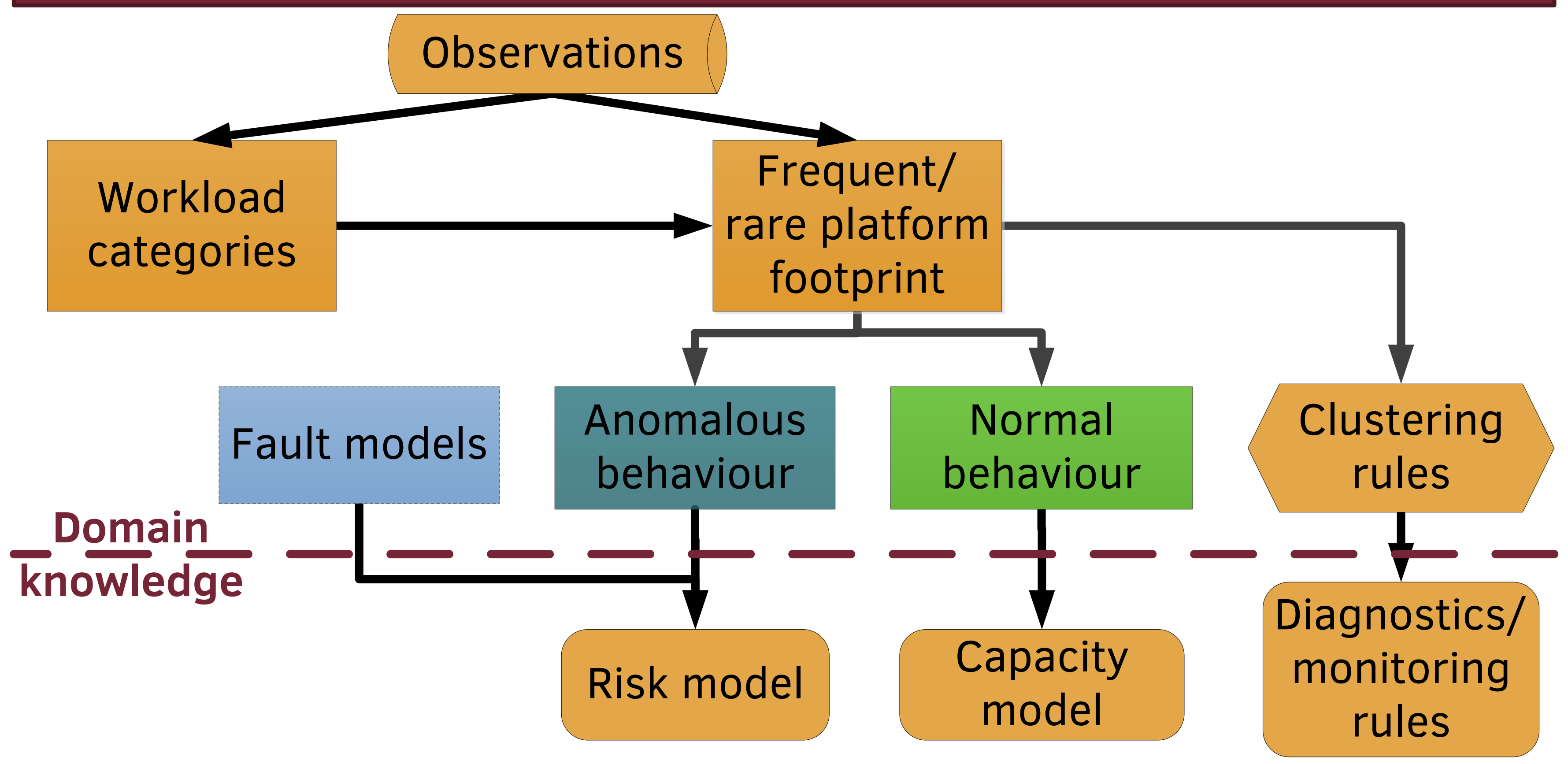
- Variability of VM charateristics → no smoothing





- High variance of network metrics → normal domain is biased
- Possible origins: stand-by VMs, system maintenance, network intrusion

Conclusion



References

[1] Pataricza et al. "Empirical Assessment of Resilience." *Software Engineering for Resilient Systems*. Springer Berlin Heidelberg, 2013. 1-16.
[2] Billor et al. "BACON: blocked adaptive computationally efficient outlier nominators." Computational Statistics & Data Analysis 34.3 (2000): 279-298.

For further information, contact:

Prof. Dr. András Pataricza
BUTE Dept. of Measurement and Information Systems
Fault Tolerant Systems Research Group

Email: pataric@mit.bme.hu
www.inf.mit.bme.hu/en/members/pataric

Ágnes Salánki, Imre Kocsis, András Pataricza