ML-based methodology for HPC facilities supervision

ISC HPC International Workshop on Monitoring and Operational Data Analytics (MODA23)

Laetitia Anton, Sophie Willemot, Sébastien Gougeaud, Soraya Zertal
Contents

1. **Context**
   Limitations & Targets

2. **Related works**

3. **Workflow**
   Phase 1: Collecting & filtering raw data
   Phase 2: Clustering & Statistical analysis
   Phase 3: Visualisation & Automation

4. **Results and discussion**

5. **Conclusion and future works**
context

- TGCC computing center
  - Joliot-Curie supercomputer: 22 Petaflops
  - Topaze supercomputer: 8.8 Petaflops

- HPC facilities
  - Infrastructure temperature (compute nodes, fans, water circuit)
  - External temperature
  - Energy consumption (computation, facilities)
  - Water consumption
  - Fan rotation

→ Administrators need to check everything
Limitations

- Huge amount of information to process
- Increase of the number of heterogenous components
- Supervision is getting complex
  - Anomalies are more difficult to detect
Targets

- Better visualisation with automation
- Detection of abnormal behaviour
- Reduction of energy consumption
Related works

Works around AI for operational HPC:

- To predict anomalies/health status on hardware
- To predict workload intensity on HPC
- To predict energy consumption of HPC facilities

➢ But at our knowledge, none for consumption tuning around facilities
Workflow
Phase 1: Collecting and filtering raw data

- **Raw data:**
  - Two years of collected data
  - Frequency of one measurement per minute
  - **93 probes** from different devices & levels of the infrastructure
Phase 1: Collecting and filtering raw data

- **Raw data:**
  - Two years of collected data
  - Frequency of one measurement per minute
  - **93 probes** from different devices & levels of the infrastructure

- **Filtered data:**
  - Unreliable and incomplete data removed: **89 probes** finally retained
  - Data reduction factor of **4.5**: for 1 year from **1.2GB to 264MB**
  - Focus on 1st quarter 2020
Phase 2: Clustering

- Target: grouping similar events into typical behaviors
- HDBSCAN as clustering method: based on cluster density, number of clusters auto-determined by setting a minimum density threshold

Power (kW) used by IT Q1 2020
Phase 2: Statistical analysis

- Identification of clusters with the same operating mode
  - Operating metrics $X$: external temperature and workload intensity
Phase 2: Statistical analysis

- Identification of clusters with the same operating mode
  - Operating metrics $X$: external temperature and workload intensity

- Determination of the deviant metrics and their impact

  Why are there 2 clusters in the same operating mode?
  - Calculation of the impact of each involved metric by measuring the gap between the quartiles of the 2 clusters
Phase 2: Statistical analysis — Application

- Identification of clusters with the same operating mode

➤ Cluster 4 and 0 share the same operating mode

---

Phase 1

Phase 2

Phase 3

Clustering

Visualization

Manual actions
Infrastructure Admin

Automation

Auto-management devices

RAW Data Files

Filtering RAW Data

Statistical Analysis

Power (kW) used by IT Q1 2020

External temperature (°C) Q1 2020
Phase 2: Statistical analysis — Application

- Determination of the deviant metrics and their impact

- Abnormal value for this inverter: an anomaly has been found
Phase 3: Automation

- Based on Dash, an open source library available in Python 3
- Cluster comparisons are done automatically
Phase 3: Visualisation

- t-SNE algorithm (t-distributed Stochastic Neighbor Embedding) used to represent multi-dimensional data in a 2D representation.

![t-SNE visualisation of the clustered dataset Q1 2020](image-url)
Phase 3: Visualisation

- t-SNE algorithm (t-distributed Stochastic Neighbor Embedding) used to represent multi-dimensional data in a 2D representation

- All these visualisations (time-series, t-SNE, 3D) with or without clustering are available to our administrators in the application

`t-SNE visualisation of the clustered dataset Q1 2020`
Results and discussion

- Capable of detect anomaly
  - Detection of an inconsistency upon power cables
Results and discussion

- Capable of detect anomaly
  - Detection of an inconsistency upon power cables

- What’s about PUE ?
  - Look for the most stable and lowest PUE
  - Apply the parameters of one cluster rather than another
Results and discussion

- Capable of detecting anomaly
  - Detection of an inconsistency upon power cables

- What’s about PUE?
  - Look for the most stable and lowest PUE
  - Apply the parameters of one cluster rather than another

→ A guideline for PUE improvement according to the workload is suggested to the infrastructure admin

![PUE variation Q1 2020](image)
Conclusion and future works

- Generic supervision methodology with ML-based algorithms
  - clear and dynamic supervision
  - highlight similar or abnormal behaviours
Conclusion and future works

- Generic supervision methodology with ML-based algorithms
  - clear and dynamic supervision
  - highlight similar or abnormal behaviours

- Tool available for infrastructure admins
  - preparation and clusterisation of data
  - visualisation of time series, t-SNE, 3D
  - automatic cluster comparisons
Conclusion and future works

- Generic supervision methodology with ML-based algorithms
  - clear and dynamic supervision
  - highlight similar or abnormal behaviours

- Tool available for infrastructure admins
  - preparation and clusterisation of data
  - visualisation of time series, t-SNE, 3D
  - automatic cluster comparisons

- Prospects:
  - implementation of alarms to warn of changes within the same operating mode
  - anticipation of deviant behaviour
Questions?

Laetitia Anton, laetitia.anton@cea.fr

Sophie Willemot, sophie.willemot@cea.fr

Sébastien Gougeaud, sebastien.gougeaud@cea.fr

Soraya Zertal, soraya.zertal@uvsq.fr