DOI 10.5162/SMSI2023/P51

Optimization of energy efficiency of an HPC cluster: On metrics, monitoring and digital twins

Alexander Kammeyer[1,2], Florian Burger [1], Daniel Lübbert [1], Katinka Wolter[2]
[1] Physikalisch-Technische Bundesanstalt, Abbestraße 2-12, 10587 Berlin, Germany,
[2] Freie Universität Berlin, Takustraße 9, 14195 Berlin, Germany
alexander.kammeyer@ptb.de

Summary:
In-depth monitoring together with a digital twin model allows to manage HPC clusters so that they can react to different external events. We shall use this information to improve operation of the cluster with respect to energy usage and cost of electricity, in order to maintain operation under reduced energy availability or to decrease the cluster's CO_2 footprint. Different additional external data sources together with system monitoring and modelling allow to adjust the HPC cluster to meet these goals.

Keywords: HPC, energy efficiency, energy supply, scheduling, digital twin, monitoring

Introduction

High Performance Computing (HPC) is the backbone of many research activities. The available compute performance has increased steadily over the last years. Despite of advances in energy efficiency of CPUs the total energy required continues to increase [1].

With the climate protection plan 2050 and climate protection program 2030 of the German federal government two initiatives exist that define goals for energy efficient data centres including the Blue Angel [2] certificate.

With energy becoming an increasingly limited resource and constant cooling becoming more and more difficult on record-breaking warm summer days, new ways of adapting an HPC cluster to such conditions are required.

The PTB operates an HPC cluster with a maximum power budget of approximately 30 kilowatts. This cluster is used by many research groups within PTB. Many other research facilities operate clusters of similar size. Lessons learned by this work can be applied to many other clusters as well.

This paper presents a new digital twin for an HPC system. It combines an in-depth monitoring system with additional external data to make cluster operation more dynamic using the digital twin and allows the system to react to events.

Efficiency metrics

Different energy metrics exist that cover different aspects of the HPC cluster. The four pillar framework [3] allows a classification of metrics and identifies which part of the data centre they cover.

Metrics, like Power Usage Effectiveness (PUE), focus on the whole data centre. Others, like Energy to Solution, focus more on the effectiveness of the hardware and software. The Blue Angel certificate uses the Energy Usage Effectiveness, a metric to be used at PTB.

Increasing effectiveness can also include reusing energy, as covered by the Energy Reuse Effectiveness metric. The Green Energy Coefficient describes how much green energy is used by the data centre. Similarly, CO_2 footprints can be quantified as described later.

Monitoring and documenting energy usage

In order to reduce energy consumption and classify energy reduction efforts, insight into the energy usage is necessary. The HPC cluster is to be fitted with a comprehensive energy monitoring. This allows to track energy usage, which is a requirement for the Blue Angel certificate and other metrics. The Blue Angel certificate requires a continuous, automatic monitoring not only of the energy but also cooling and system usage.

The monitoring can also be used to adjust the energy consumption, so that the consumption remains within defined bounds.

Digital twin

The operation of the HPC cluster is influenced by the operating system, scheduler and resource manager. A digital twin of the cluster helps these components to make predictions about the cluster in the future. This includes the available energy and cost of the energy, expected energy consumptions of jobs and thermal load.

One possible use case is Power Aware Scheduling. With power aware scheduling, the system can be tuned to reduce overall power consumption [4]. Unused nodes can be turned off [1].

Demand Response

Demand response refers to a technique usually found in heavy industry. With a mutual agreement between the energy service provider (ESP) and consumer, the ESP notifies a consumer of an upcoming energy shortage and requests to lower the energy consumption. In return, the consumer receives a financial benefit or can remain operational as opposed to being cut from energy entirely.

This method can be adopted for HPC [5]. An HPC cluster offers many ways of reducing energy consumption. Possible ways are suspending running jobs, disabling idle compute nodes or reducing CPU frequency on the entire cluster. Even temporarily altering the scheduling behaviour, e.g. preventing new jobs from running, are conceivable. The digital twin model of the cluster helps to correctly predict the impact of the different methods and make the optimal decision.

CO_2 footprints

The energy in the energy grid consists of a mix of different energy sources like solar, wind and water but also gas, coal and nuclear energy. Each source has a different impact on the climate, most notably in terms of CO_2 emissions.

The current energy mix can be taken into account and allows to estimate the CO_2 footprint for the cluster. Jobs can be scheduled with the goal of reducing emissions and with the help of energy monitoring emissions can be tracked.

Cost of electricity

With energy becoming a more limited resource, companies might be required to buy energy packets on the so-called spot market to daily or even hourly changing rates.

If that is the case, it has a direct influence on the operating cost of the HPC cluster. This could be used to optimize the cost or make sure, that certain levels are not exceeded in times of high market prices. Jobs can be placed in times of lower prices or operation can be suspended if energy costs rise above a certain threshold.

Influence of weather conditions

The weather has an influence on the energy efficiency metrics because it directly influences the required energy for cooling. The metrics are therefore often averages over a year.

With the adaptation of a free cooling system, cooling becomes more efficient compared to a classic cooling machine but might not be feasible on extremely warm days. The HPC cluster needs to respond to such events of limited cooling. Using the comprehensive monitoring together with weather forecast data, the system scheduling can reduce the system load on predicted warm days when the outside temperature comes close to operational limits.

Summary

With the introduction of a comprehensive monitoring and the selection of metrics for energy efficiency, the energy optimization efforts for the HPC cluster at PTB can be measured and quantified.

With a digital twin for the HPC cluster, optimization efforts become easier as it combines different external data sources with the cluster monitoring. This combination allows the real system to adapt to different scenarios.

The monitoring and digital twin play an instrumental role when optimizing energy consumption, scheduling jobs during times of lower energy costs or outside extreme heat periods and help to operate the cluster within defined limits.

References

[1] Mämmelä, Olli et al., Energy-aware job scheduler for high-performance computing, *Computer Science - Research and Development 27*, 265-275 (2012); doi: 10.1007/s00450-011-0189-6

[2] RAL UMWELT, Energieeffizienter Rechenzentrumsbetrieb DE-UZ 161, 2 edition (2015)

[3] Wilde, Torsten et al., The 4 Pillar Framework for energy efficient HPC data centers, *Computer Science - Research and Development* 29, 241-251 (2014); doi: 10.1007/s00450-013-0244-6

[4] Bodas, Deva et al., Simple Power-Aware Scheduler to Limit Power Consumption by HPC System within a Budget, *2014 Energy Efficient Supercomputing Workshop*, 21-30 (2014); doi: 10.1109/E2SC.2014.8

[5] Ahmed, Kishwar et al., Enabling Demand Response for HPC systems through Power Capping and Node Scaling, *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 789-796 (2018); doi: 10.1109/HPCC/SmartCity/DSS.2018.00133