HPI Data Center Climate Footprint II

(Bachelor project, Winter 2023) - Data Engineering Systems Group

Motivation

Computer science has long been seen as a neutral actor relation to climate change. However, data centers and telecommunication systems are using increasing amounts of energy. As a prominent example, the BitCoin network alone is estimated to consume as much energy as many countries [1].

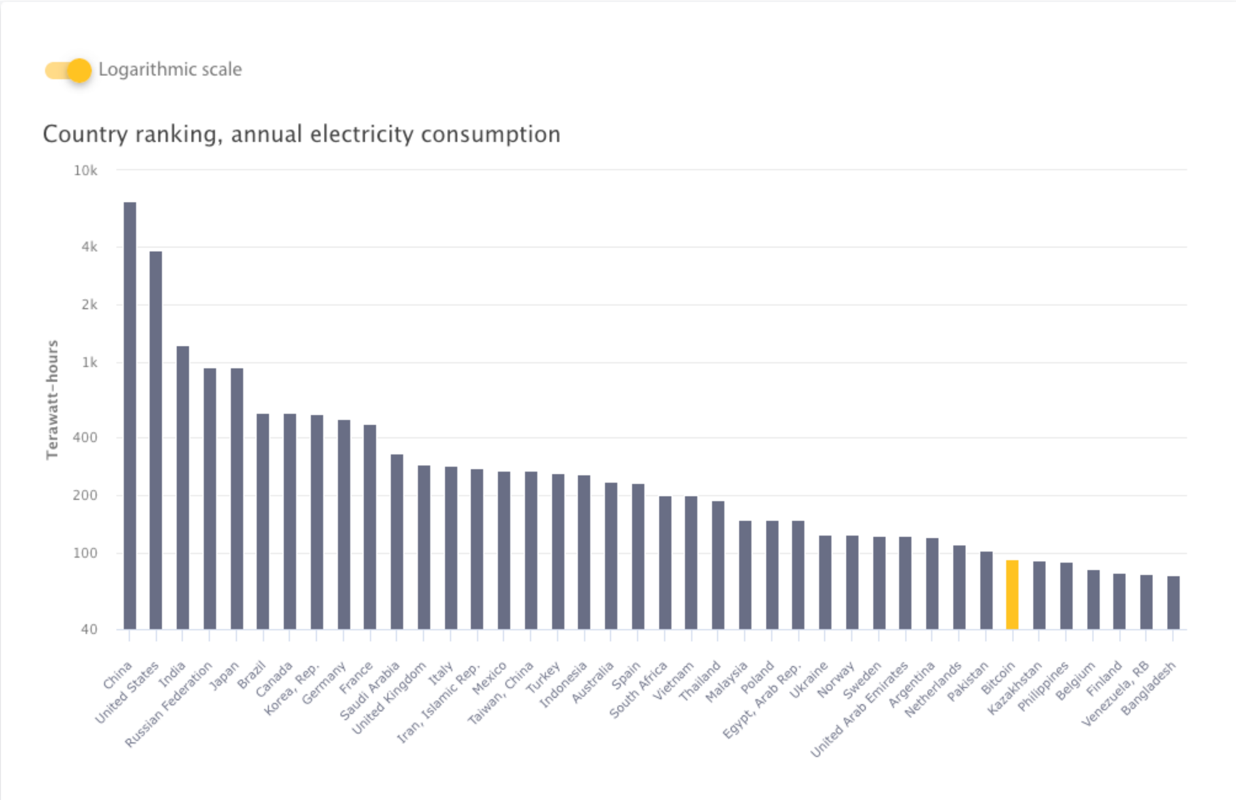

Figure 1 Country comparison - Source: https://ccaf.io/cbeci/index/comparisons

While energy consumption is a prominent factor in climate footprint, it is not the only factor. Server production and shipping also have a high impact that is often neglected [2]. Economically, energy consumption of a server is not the major driver of data center cost and, therefore, data centers aim for constant (efficient) power consumption rather than optimizing for power savings [3]. Because of this, many research projects in energy efficiency in data centers have limited practical applicability.

With net zero targets in the EU for 2050, it becomes even more important to holistically measure the climate footprint of IT based systems and services, in order to fully understand the driving factors of emissions. Only with this knowledge, it is possible to efficiently reduce the climate footprint and offset remaining emissions.

With the new data center at HPI, we have a real example of a data center with all infrastructure. Goal of the project is to build a model and framework to calculate the climate footprint of our data center, cradle to cradle. This means, we want to incorporate all sources of greenhouse gases in the model (Scope 3 in the Greenhouse Gas Protocol [5]).

Project description

Building on the previous bachelor project, the main objective of this project is to build a model and platform for collecting and estimating information of the climate footprint of the HPI Data Center and mapping it to compute jobs and services in the data center. This will include operational emissions through energy consumption as well as emissions from transport, construction, production, travel, etc. The input data will be stored in a database and the model needs to be adjustable to include new factors that could influence the footprint (e.g., adding solar panels, reusing server heat, or new evaluation of emissions). The previous bachelor project has built a dashboard to present the complete data center CO2 footprint. In this project, we want to further refine the model and map the emissions to the utilisation in the data center.

Current view of the data center dashboard.

The project will be built in several phases, starting with breakdowns per rack and server, starting with the direct energy consumption. The framework should then produce overviews of the daily, monthly, yearly consumption and climate footprint in a dynamic dashboard. Furthermore, we want to create updateable reports for compute jobs in the data center to make users aware about the power consumption and climate footprint.

The model and framework will be generic to be reusable for other data centers.

Project partners

During this project, the participants will collaborate with colleagues from Hewlett Packard Enterprise and climate researcher Stefan Krottenthaler from University of Passau.

Skills

The participants need to have experience in software engineering and at least one programming language (C++, Java, Python). Basic experience in machine learning is beneficial. Participants should be comfortable with documenting their work and visualizing their results. The initial phase of the project will include literature research to find models and estimations for emission calculation.

During the project, you will further your knowledge about software engineering in teams and programming, but also get deep insights into data center architecture and setup as well as climate impact prediction and analysis.

Contact

Please contact Tilmann Rabl tilmann.rabl@hpi.de for any questions.

Recommended reading

1. https://ccaf.io/cbeci/index/comparisons
2. https://www.goclimate.com/blog/the-carbon-footprint-of-servers/
3. In Computer Architecture, We Don't Change the Questions, We Change the Answers, Mark D. Hill, Keynote at Data Management on New Hardware (DaMoN) Workshop @ SIGMOD, June 2022. Slides: pdf
4. https://ec.europa.eu/clima/eu-action/climate-strategies-targets/2050-long-term-strategy_en
5. https://de.wikipedia.org/wiki/GHG_Protocol