Towards low-carbon globally distributed clouds

Green Days 2023 @ Lyon

Miguel SILVA VASCONCELOS[1,2]

Mars 2023

School of Sciences, Arts, and Humanities, University of São Paulo, Brazil[1]
Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG, Grenoble, France[2]

The energy needed for Cloud

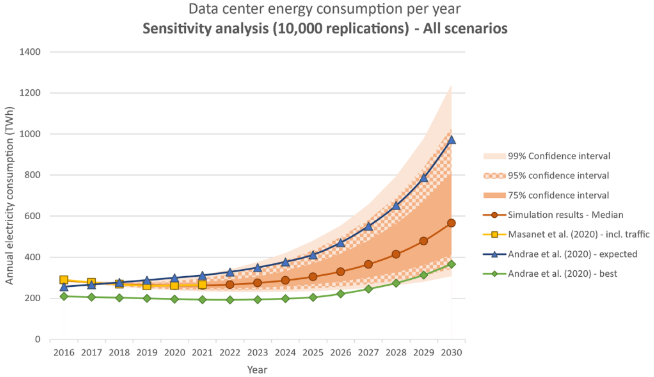

Figure 1: Koot and Wijnhoven, “Usage impact on data center electricity needs: A system dynamic forecasting model.” Applied Energy, 2021.

Reducing the environmental impact of cloud operations

- Projects already deployed (or in development) by major cloud providers (Amazon AWS, Apple, Facebook, Google, Microsoft)

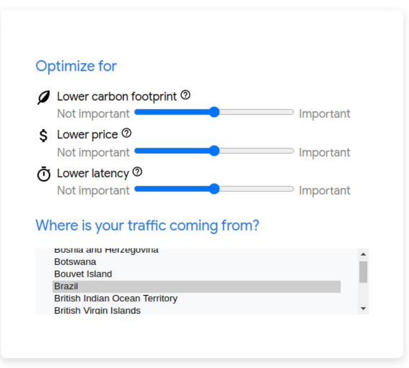

Figure 2: Google's region picker tool.

Renewable energy to reduce the environmental impact of the clouds

Intermittent nature of renewables

- *time of the day*, *weather*, and *season of the year*
- Solar power:
  - Costs have fallen by 85% from 2010 to 2019
  - Solar irradiation has lower variation than wind speed
- Lithium-Ion batteries:
  - Efficiency in terms of costs, power and energy density, charge and discharge ratio and self-discharge

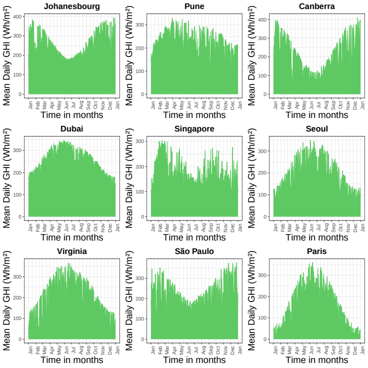

Figure 3: Solar Irradiation at different locations in 2021.

Carbon-Responsive Computing

Approaches that are aware of their carbon intensity (carbon aware) and make informed decisions[1]

- Follow-the-renewables
- Sizing the DCs renewable infrastructure

[1]Dawn Nafus, Eve M. Schooler, and Karly Ann Burch. "Carbon-Responsive Computing: Changing the Nexus between Energy and Computing." In: *Energies* 14.21 (2021). ISSN: 1996-1073. DOI: 10.3390/en14216917. URL: https://www.mdpi.com/1996-1073/14/21/6917.

Carbon-Responsive Computing

Approaches that are aware of their carbon intensity (carbon aware) and make informed decisions[2]

- Follow-the-renewables
  - Allocates/Migrates the workload to the data centers (DCs) that have more renewable (green) power available
  - Migrating the workload among different DCs generates extra computations proportional to the duration of the migration
- Sizing the DCs renewable infrastructure

[2] Dawn Nafus, Eve M. Schooler, and Karly Ann Burch. "Carbon-Responsive Computing: Changing the Nexus between Energy and Computing." In: *Energies* 14.21 (2021). ISSN: 1996-1073. DOI: 10.3390/en14216917. URL: https://www.mdpi.com/1996-1073/14/21/6917.

Impact of follow-the-renewables[3]

- Baselines that do not consider network:
  - Migration time > 100 times longer
  - Wasted energy could have powered one of the DCs for 44 hours
- Proposed solution:
  - Migration algorithm that considers network bandwidth, topology, and the history of usage of the links
  - No network congestion and same or lower brown energy consumption

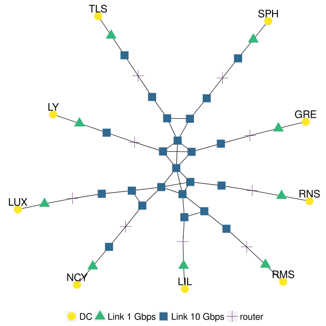

Figure 4: DCs and how they are connected in the network.

[3] Miguel Felipe Silva Vasconcelos, Daniel Cordeiro, and Fanny Dufossé. "Indirect Network Impact on the Energy Consumption in Multi-clouds for Follow-the-renewables Approaches." In: *11th International Conference on Smart Cities and Green ICT Systems*. 2022.

Approaches that are aware of their carbon intensity (carbon aware) and make informed decisions[4]

- Follow-the-renewables
- Sizing the DCs renewable infrastructure
    - Compute the area of solar panels (PVs) in m^2 and batteries capacity in Wh

[4] Dawn Nafus, Eve M. Schooler, and Karly Ann Burch. "Carbon-Responsive Computing: Changing the Nexus between Energy and Computing." In: *Energies* 14.21 (2021). ISSN: 1996-1073. DOI: 10.3390/en14216917. URL: https://www.mdpi.com/1996-1073/14/21/6917.

Data centers:

- Infrastructure already built (servers, network)
- Homogeneous (regarding CPU cores)
- Server power consumption: idle and dynamic
- Intra network power consumption: static
- Specific Power Usage Effectiveness (PUE) for each DC

Figure 5: Selected data centers location (inspired from Microsoft Azure)

Sizing renewable infrastructures

Workload:

- All tasks must be scheduled and executed on time
- Batch tasks that can be executed in any of the DCs
- No migration
- Task execution cannot be delayed

Sizing renewable infrastructures

Renewables infrastructure

- Batteries charge and discharge efficiency, Maximum Depth of Discharge
- PV panels efficiency
- Carbon emissions from manufacturing (PV: 250 kg CO_2 eq per m², bat: 59 kg CO_2 eq per kWh)
- Lifetime (PV: 30 years, bat: 10 years)

Sizing renewable infrastructures

Local electricity grid

- The energy mix is different at each location
- May have the presence of renewables or low carbon-intensive sources

Table 1: Emissions (in g CO_2-eq/kWh) for using the regular grid. Source for grid emissions: electricityMap, climate-transparency.org.

Location	Emissions
Johannesburg	900.6
Pune	702.8
Canberra	667.0
Dubai	530.0
Singapore	495.0
Seoul	415.6
Virginia	342.8
São Paulo	61.7
Paris	52.6

Proposed solution

Linear program formulation to minimize the carbon emissions from the cloud federation operation (timespan of 1 year) [2]

- Scheduling and dimensioning modeled as single problem
  - Allocate workload to other DC or increase the battery capacity or PV area?
- Only real variables
  - Optimal solution in polynomial time: 394264 variables, solved in less than 1 minute with Gurobi

[2] M. Vasconcelos, D. Cordeiro, G. Da Costa, F. Dufossé, J.-M. Nicod, and V. Rehn-Sonigo, "Optimal sizing of a globally distributed low carbon cloud federation". *In: 2023 23nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid), Bengaluru, India, 2023.*

LP Overview

Data center power consumption:

$$P_k^d \leq Pre_k^d + Pgrid_k^d + Pdch_k^d - Pch_k^d \tag{1}$$

where Pch_k^d is the power to charge the battery at each time of time slot k on DC^d and $Pdch_k^d$ is the power to discharge the battery, Pre_k^d is the solar power produced, and $Pgrid_k^d$ is the power used from the local grid.

Workload:

$$w_k^d \leq C^d \tag{2}$$

where w_k^d is number of cores needed during the kth time slot on DC^d, and C^d is the number of cores within DC^d .

Batteries level of energy (B_k^d):

$$B_k^d = B_{k-1}^d + Pch_{k-1}^d \times \eta_{ch} \times \Delta t - \frac{Pdch_{k-1}^d}{\eta_{dch}} \times \Delta t \qquad (3)$$

where η_{ch} is efficiency of the charge process and η_{dch} is the efficiency of the discharge process.

Solar power production:

$$Pre_k^d = I_k^d \times Apv^d \times \eta_{pv} \qquad (4)$$

where I_k^d is the solar irradiance, Apv^d the PV panel area, and η_{pv} is the efficiency of PV module

LP Overview

Objective function:

$$\text{minimize} \sum_{k=0}^{K-1} \sum_{d=1}^{D} (FPgrid_k^d + FPpv_k^d) + \sum_{d=1}^{D} FPbat^d \tag{5}$$

LP Overview

Inputs:

- Real solar irradiation data from 2021 (MERRA-2)
- Real values from carbon emissions of using local electricity grid, manufacturing PVs and batteries
- Efficiency parameters of PVs and batteries
- Number of DCs, and cores inside each DC
- Power consumption of servers and network devices (Grid'5000 Taurus, HP ProCurve 2810-48G)
- PUE of each DC
- Workload (based on Google traces)

Ouptut

- Area of PVs (m^2) and capacity of batteries (Wh)

Results

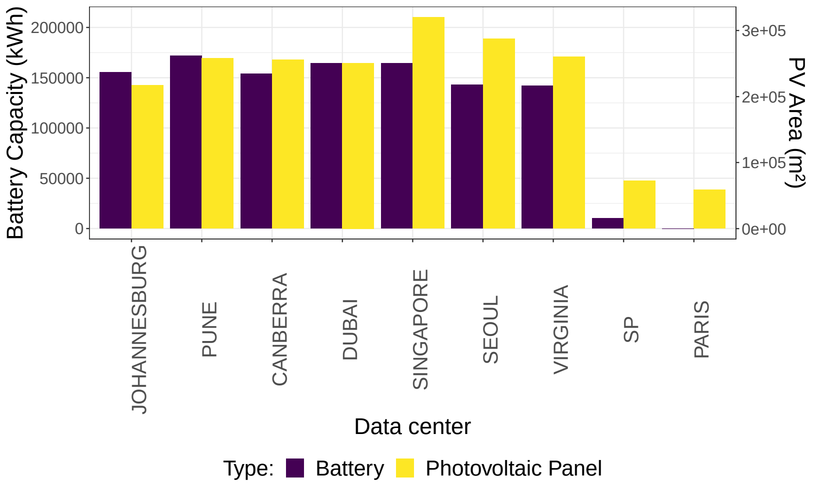

Figure 6: Optimal result for the area of PV panels and capacity of the batteries.

Results

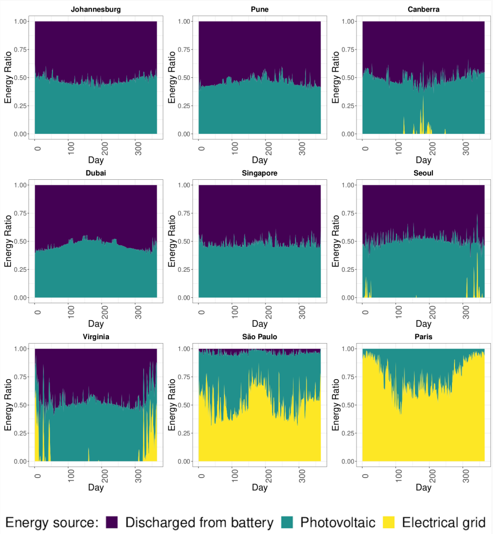

Figure 7: Composition of the DCs' daily energy consumption throughout the year considering the different sources of energy

Results

Table 2: Total emissions for the different scenarios.

Scenarios	Emissions (t CO_2-eq)
Electrical grid	201211.3
PV and batteries	42370.6
PV, batteries, and grid	29600.6

Results

Table 3: Evaluating sizing considering solar irradiation data of different years (2018, 2019, 2020) using the MAPE metric (values are in %)

Location	PV Area	Battery Capacity
Johannesburg	1.72	1.64
Pune	3.72	0.76
Canberra	8.62	4.25
Dubai	2.31	2.88
Singapore	7.22	0.34
Seoul	3.15	1.11
Virginia	2.2	0.87
São Paulo	5.81	8.05
Paris	2.76	0

Visualizating data center operation and energy source used

- Each circle is a DC, and its radius is the power consumption (in MWh)
- The pizza graph represents the share of electricity source being used at that instant (from pv panels, from the batteries, or from the grid)
- The gray shadow represents the night
- Visualization for the first week of 2021

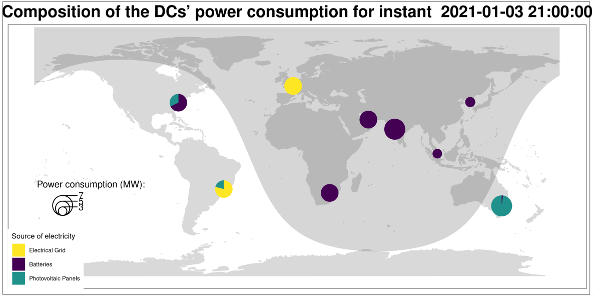

Figure 8: Example of data centers electricity source used.

Results

Example of DCs electricity source used.

Future work

- Robustness of the model
- Consider VM migration of the tasks and other scheduling algorithms
    - Quality of Service
- Other workload types (not only batch tasks)
- Dimensioning of IT infrastructure (servers considering the footprint of manufacturing, new generations that are more efficient)

Sponsor acknowledgments

This work has been partially supported by the LabEx PERSYVAL-Lab ("ANR-11-LABX-0025-01") funded by the French program *Investissement d'avenir*, by grant #2021/06867-2, São Paulo Research Foundation (FAPESP), by the EIPHI Graduate school (contract "ANR-17-EURE-0002"), by the EuroHPC EU Regale project (g.a. 956560), and by the ANR DATAZERO2 (contract "ANR-19-CE25-0016") project.

Thank you!

Towards low-carbon globally distributed clouds

Figure 9: Supplemental material for the paper: M. Vasconcelos, D. Cordeiro, G. Da Costa, F. Dufossé, J.-M. Nicod, and V. Rehn-Sonigo, “Optimal sizing of a globally distributed low carbon cloud federation”. *In: 2023 23nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid), Bengaluru, India, 2023.*

Contact:

miguel.silva-vasconcelos@inria.fr

References

Koot, Martijn and Fons Wijnhoven. "Usage impact on data center electricity needs: A system dynamic forecasting model." In: *Applied Energy* 291 (2021), p. 116798. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2021.116798.

Nafus, Dawn, Eve M. Schooler, and Karly Ann Burch. "Carbon-Responsive Computing: Changing the Nexus between Energy and Computing." In: *Energies* 14.21 (2021). ISSN: 1996-1073. DOI: 10.3390/en14216917. URL: https://www.mdpi.com/1996-1073/14/21/6917.

Vasconcelos, Miguel Felipe Silva, Daniel Cordeiro, and Fanny Dufossé. "Indirect Network Impact on the Energy Consumption in Multi-clouds for Follow-the-renewables Approaches." In: *11th International Conference on Smart Cities and Green ICT Systems*. 2022.