

From Energy Efficient Cloud Infrastructure to Energy Aware Cloud Architecture!



Cloud Service Providers (CSP's) have become increasingly aware about the climate impact of cloud computing and as a result are proposing measures to make their platforms more sustainable. However, when analyzing their approaches to sustainability in more detail, often sustainability is reduced to three aspects:

1. Optimizing the energy consumption of the infrastructure: reducing the overall energy consumption
2. Using energy from sustainable sources: using green or carbon neutral energy sources
3. Using equipment that has a sustainable life-cycle: using a green device supply chain

The energy consumption optimization is translated by all hyperscale cloud providers into:

- IT Operational efficiency
- IT Equipment efficiency
- Datacenter Infrastructure efficiency

Mapping this to cloud service models, it becomes clear that the approaches are working bottom-up in the

architecture stack, seem not to pay attention at the application and data level and focus more on infrastructure asset management.

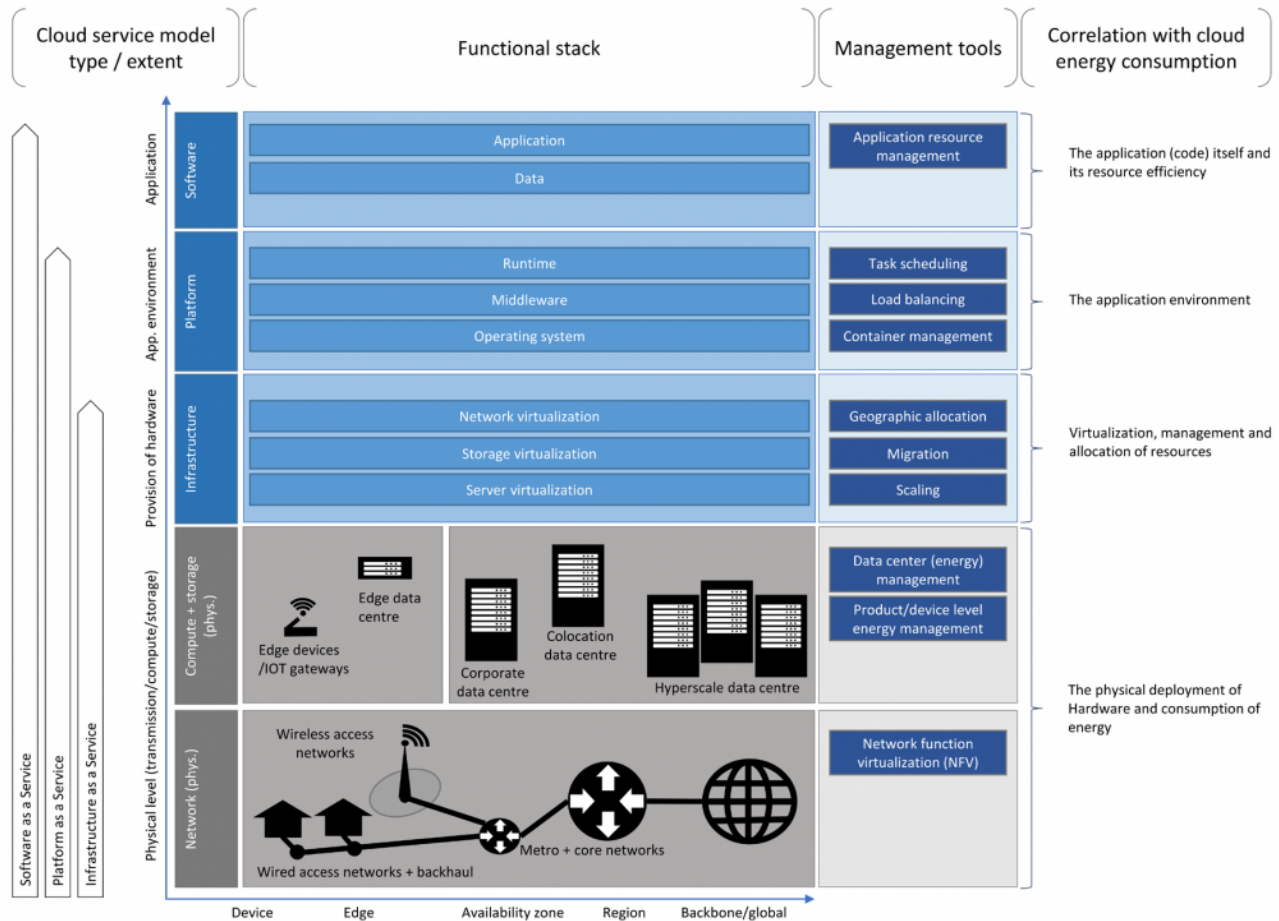
laas	Paas	Saas
Applications	Applications	Applications
Data	Data	Data
Runtime	Runtime	Runtime
Middleware	Middleware	Middleware
OS	OS	OS
Virtualization	Virtualization	Virtualization
Servers	Servers	Servers
Storage	Storage	Storage
Networking	Networking	Networking

Managed by Consumer

Managed by Vendor

It is nice to see CSP's looking at more efficient ways to organize their infrastructure but the usage of their cloud infrastructure seems to be underexposed. Efficient usage is one, but reducing the need for consumption is another. CSP's can't be blamed for our consumption but the focus on energy alone, although important, seems like green washing: *"What is not consumed does not need to be optimized or turned in something carbon neutral!"*

All this points to the fact that the approaches to cloud native sustainability is still in its infancy. Bottom-up optimization focusses on efficiency and is becoming an established approach.



EC, 2020, Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market.

In the bottom-up approach an element that often is overlooked, is the cost of organizing the energy optimization. It requires additional monitoring and operations management infrastructure which also result in cloud computing and storage requirements i.e. components that also consume energy. None of the CSP's mentions the management overhead costs of the operations management infrastructure.

The top-down optimization, focusing on the cloud usage, is still in research and wants to tune cloud efficacy in stead of efficiency. The top-down approach involves:

- Application and Data architecture
- Application Usage

Application and data architecture receives lately a bit more attention.



Firstly through the use of the **cloud-native computing paradigm**, *selecting of the right component for the right job*, has a positive effect. For example the use of containerized micro-services reduced the need for full fledged VM's what resulted in an infrastructure optimization: less compute and less storage. Important not to overlook is that this approach also introduces more network communication and complex management but the overall overhead is lower than the resource efficiency gains. Data traffic optimization can be found under **Fog Computing techniques** that are being researched.



Secondly there is the **programming efficiency** i.e. *the correct organization of algorithms* in application code. Inefficient algorithms will require more CPU cycles, memory and more storage capacity. Static code analysis tools are currently used to check for architecture qualities in code like maintainability. For the moment checking algorithm efficiency is not commonly available in such a tools and are part of specialized tooling. Programming language and compiler efficiency checks are not mainstream at all. These would involve checking the CPU instruction sets being used in the binary code. CPU instructions have different effect on the CPU's energy consumption. These techniques are being developed under the umbrella of the **Energy Aware Software Engineering**.



Application Usage has not received any attention!

Finally we have to point out to the **acceptable use of technology**. Lately a lot of focus has gone to the ethics of data privacy. As applications are running on infrastructure that are consuming natural resources, one should wonder if a cost-benefit analysis of the usage of application should not be considered: *"It is not because we can that we should use an application and in extension the cloud"*.



This opens the discussion to a new debate: **from less energy consumption towards the correct application of the available energy.** Should we start taxing applications like we do with other consumer products based on necessity? Alcohol (pleasure) is taxed more than water (basic need) could be translated into the tax on a Social Media application (pleasure) vs. Banking application (basic need)!

The following document is on Microsoft Efforts for cloud sustainability:

[Microsoft_Cloud_Carbon_Study_2018Download](#)

The EC released following report on an Eco-Friendly Cloud Market:

[final_report_as_updated_by_op_kk0320210enn_5E14973E-9C30-1BFD-306046B2C76EAAF1_71330Download](#)