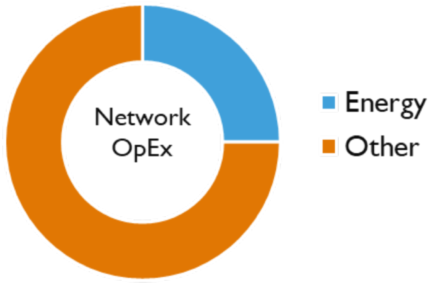
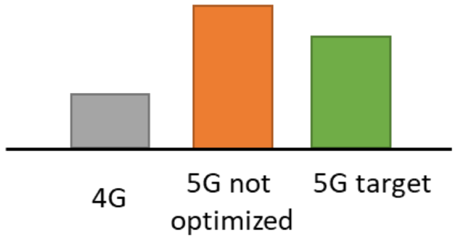
Towards a sustainable 6G: research challenges and opportunities

Raffaele Bruno
Senior Researcher
IIT-CNR, Italy
raffaele.bruno@iit.cnr.it

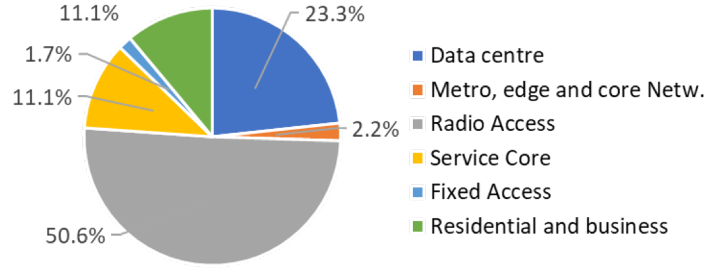
Why is energy so important for MNO?



Contribution of energy consumption in the OpEx of MNOs [1]



Energy consumption levels of MNO facilities [1]



Breakdown of the energy consumed by the network segment [2]

- The telecoms industry is responsible for 2-3% of the total power consumption of humankind.
- Energy costs today represent between 20% and 40% of a telecoms company's OPEX
- 5G can be up to 90% more energy-efficient than 4G. However,
  - huge increases in density and traffic are expected to negate these savings, leading to a net 5G energy consumption potentially 4 to 5 times higher than 4G
  - edge computing infrastructure embedded into 5G will play a key role in total energy consumption
  - Impact of network softwarisation on the 5G energy efficiency may be huge

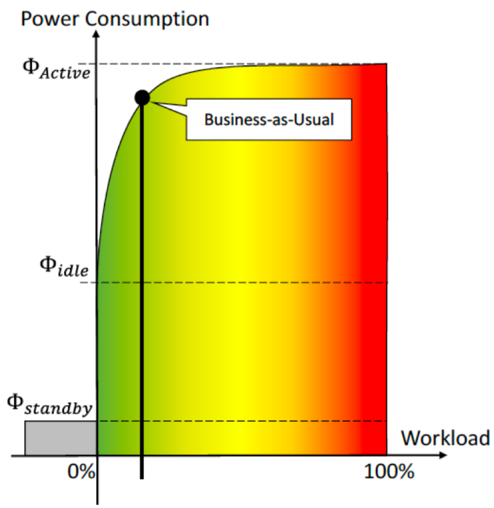
[1] "Energy Efficiency: an Overview," GSMA Future Networks, 8th May 2019. URL: https://www.gsma.com/futurenetworks/wiki/energy-efficiency-2/.
[2] Lorincz J, Capone A, Wu J. Greener, "Energy-Efficient and Sustainable Networks: State-Of-The-Art and New Trends." Sensors. 2019; vol. 19, no. 22:4864. https://doi.org/10.3390/s19224864

Energy efficiency (EE) from physical to virtualised networks

- Historically, energy efficiency in networking technologies has been commonly interpreted as the ratio between the network capacity and its energy consumption.
- Approaches for increasing energy efficiency:
  1. Bandwidth overprovisioning (i.e., planning for the peak demand)
  2. Energy-efficient hardware design (e.g., low-power modes)
- Physical Network Functions (PNFs) are still tightly coupled with the underlying hardware substrate
  - Focus is still on energy-efficient hardware

$$EE = \frac{Bandwidth}{Energy\ Consumption}\ [b/J]$$

- Software virtualisation technologies (e.g. NFV) are typically considered intrinsically green because on-demand allocation of VNFs allows horizontal scaling according to the to traffic. However,
  - higher utilisation of server resources may lead to higher power consumption
  - general-purpose hardware is not energy efficient
    - virtualized BBU causes consumptions around 250% higher with respect to the commercial deployment, and operation and capital costs over 66% higher [1]

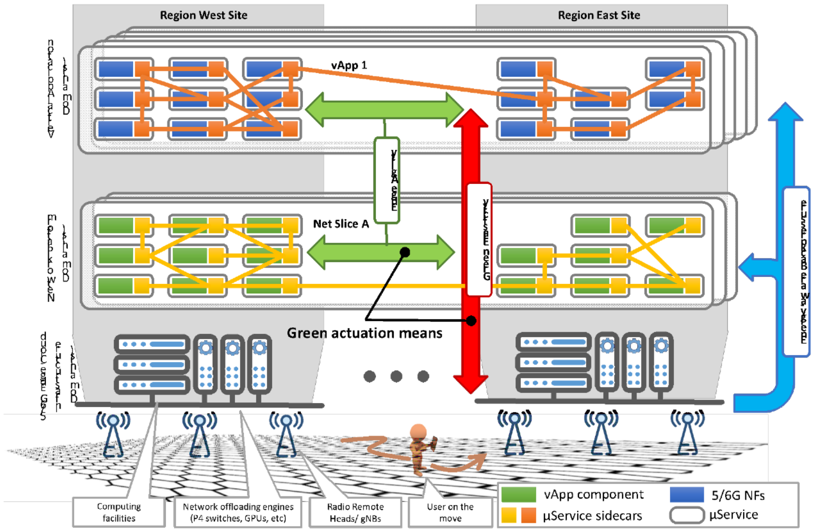


[1] R. Bolla, R. Bruschi, F. Davoli, C. Lombardo and J. F. Pajo, "Debunking the "Green" NFV Myth: An Assessment of the Virtualization Sustainability in Radio Access Networks," 2020 6th IEEE Conference on Network Softwarization (NetSoft), Ghent, Belgium, 2020, pp. 180-184,

The 6Green project (https://www.6green.eu/)

- Project at a glance:
  - Duration : 1st January 2023: - 31st December 2025
  - Call Topic ID: HORIZON-JU-SNS-2022-STREAM-A-01-04
  - Consortium: 15 Partners
  - Requested EU Contribution: € 5,996,896.00
  - Project Coordinator: CNIT – Prof. Roberto Bruschi

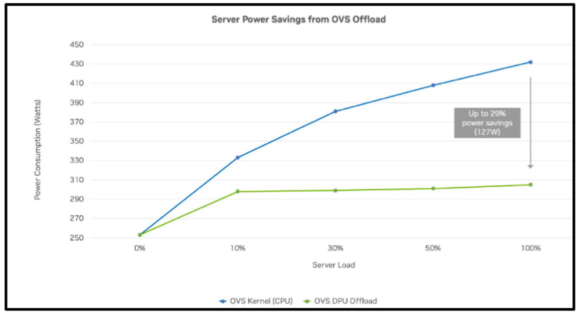
- Objective? promote energy efficiency across the whole 6G value-chain, to enable both 6G networks and vertical applications to reduce their carbon footprint
- How? extend cloud-native technologies and the B5G SBA with new cross-domain technology enablers:
  - Green elasticity
  - Edge Agility
  - Energy-aware observability

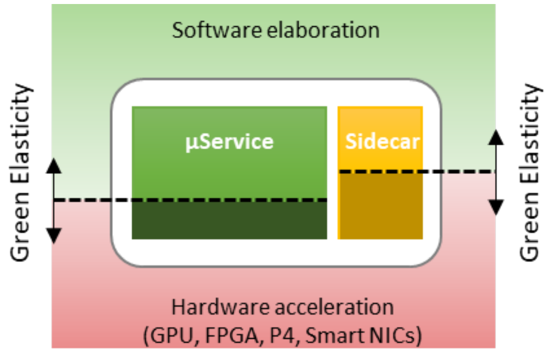


The green elasticity

- A variety of hardware accelerators (GPUs, eBPF, DPDK, DPU) are available that include purpose-built processors for accelerating a specific data processing function or workload

- hardware acceleration engines exhibit low power-consumption dependency against their usage (example from an Open vSwitch (OVS) networking offload to the to the networking accelerator on the BlueField DPU [1])





- <u>Solutions for</u> providing energy-aware hardware-assisted acceleration to network functions and vertical applications to enable smart vertical scalability with the twofold objective of
    1. distributing end-to-end latency budgets of vertical applications across the domains,
    2. optimizing the trade-off between the energy/carbon footprint and the performance of network and application artefacts.

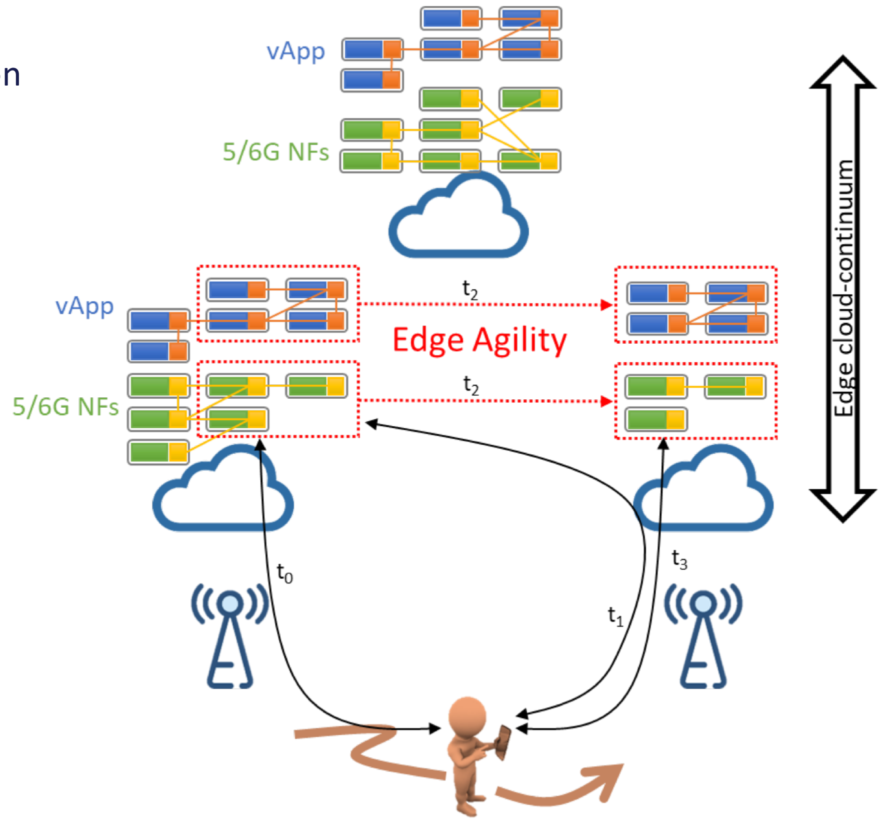
[1] NVIDIA White Paper, "DPU Energy Efficiency", November 2022

The edge agility

- The edge agility is a a sort of handover procedure for SBA/application side that provide horizontal scalability to vertical applications and related network slices across the 5/6G network-edge-cloud continuum

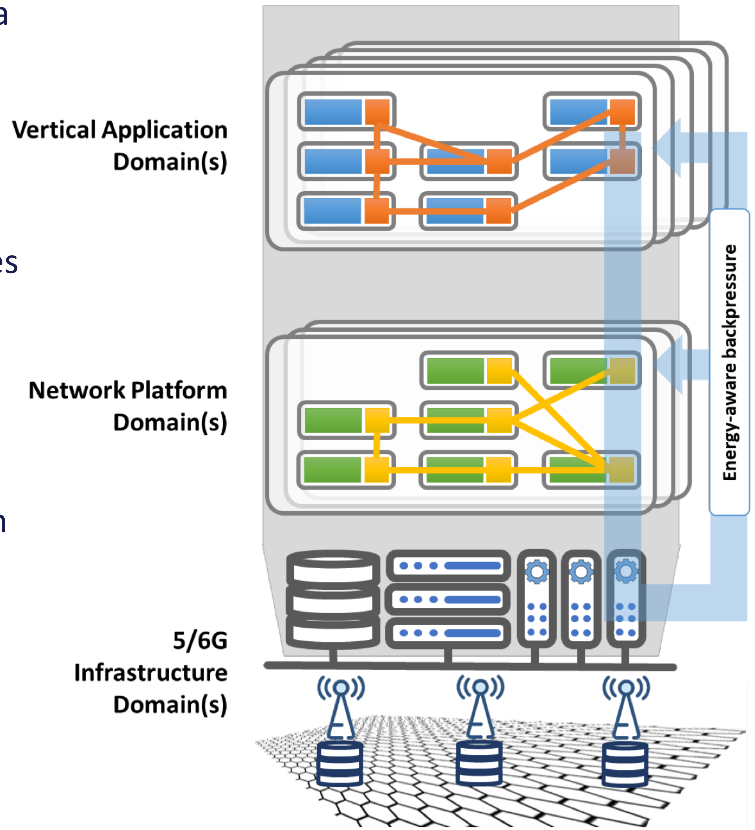
- Workload redistribution according to user or infrastructure-driven events, (e.g., user mobility, seamless workload replacement/migrations, etc.). Opportunistically move the latency budget between connectivity and computing

- Edge agility should permit to rapidly "scale to zero" the footprint of the slice/vertical application in all not used continuum areas, and to quickly resume the operating capacity when needed



Energy-aware observability

- Energy accounting is crucial for optimising power provisioning. However, in a NFVI the finest grain at which energy consumption can be measured is the NFVI Node, making it almost impossible to measure the energy consumed by VNFs separately
- Partial solutions exist for data centres, e.g., providing power models for estimating the power consumption at the process, container, and Kubernetes pod level
- Energy-aware backpressure is a set of cross-domain observability mechanisms and analytics to evaluate the energy consumption on a per slice/vertical application level
  - Hardware-level energy consumption metrics can be mapped onto each hosted tenant through AI-driven analytics
- From SLAs to DLAs (Decarbonisation Level Agreements)



Take-away messages

- Network softwarisation is transforming the way we plan and operate mobile networks
- RAN is still the main source of power consumption on 5G network, but 5GC contribution is around 25%
- To optimise energy consumption, we should leverage the availability of heterogenous hardware accelerators
- Dynamic reconfiguration/migration of slices/application to adjust to workload distribution
- Energy accountability on a per tenant/slice/vertical application level is crucial for energy optimisation