

Greening Networking: Toward a Net Zero Internet

Alexander Clemm^{*1}, Dirk Kutscher^{*2}, Michael Welzl^{*3},
Cedric Westphal^{*4}, Noa Zilberman^{*5}, and Simone Ferlin-Reiter^{†6}

1 Independent – Los Gatos, CA, US. ludwig@clemm.org

2 HKUST – Guangzhou, CN. dku@hkust-gz.edu.cn

3 University of Oslo, NO. michawe@ifi.uio.no

4 University of California, Santa Cruz, US. cedric@ucsc.edu

5 University of Oxford, GB. noa.zilberman@eng.ox.ac.uk

6 Red Hat – Stockholm, SE & Karlstad University, SE. simoneferlin@gmail.com

Abstract

This is the report of the Dagstuhl Seminar 24402 on *Greening Networking: Toward a Net Zero Internet* that took place from September 29th to October 2nd 2024. The seminar discussed the most impactful networking improvements for reducing carbon emissions in three different areas: 1) applications, systems, and stakeholders; 2) network technologies; and 3) lifecycle and control loops. As a major result of the seminar, the following problems and topics for future research were identified: 1) characterizing the Internet footprint on carbon emissions accurately; 2) understanding attributional and consequential accounting of carbon emissions in networked systems; and 3) identifying potential solutions to give network systems more flexibility in better supporting energy grids and connecting to renewable energy sources. One of the concrete results of this seminar is a list of technologies and research opportunities for which we estimated the potential impact and time horizons.

Seminar September 29 – October 2, 2024 – <https://www.dagstuhl.de/24402>

2012 ACM Subject Classification Networks; Networks → Network management; Networks → Network protocol design; Networks → Network algorithms; Networks → Network dynamics; Hardware → Power and energy

Keywords and phrases Green networking, carbon-aware networking, sustainable networking, energy-efficient networking

Digital Object Identifier 10.4230/DagRep.14.9.167

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Greening Networking: Toward a Net Zero Internet, *Dagstuhl Reports*, Vol. 14, Issue 9, pp. 167–192

Editors: Alexander Clemm, Dirk Kutscher, Michael Welzl, Cedric Westphal, and Noa Zilberman



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary


Alexander Clemm (*Independent – Los Gatos, US, ludwig@clemm.org*)

Dirk Kutscher (*HKUST – Guangzhou, CN, dku@hkust-gz.edu.cn*)

Michael Welzl (*University of Oslo, NO, michawe@ifi.uio.no*)

Cedric Westphal (*University of California, Santa Cruz, US, cedric@ucsc.edu*)

Noa Zilberman (*University of Oxford, GB, noa.zilberman@eng.ox.ac.uk*)

License  Creative Commons BY 4.0 International license

© Alexander Clemm, Dirk Kutscher, Michael Welzl, Cedric Westphal, and Noa Zilberman

The Dagstuhl Seminar focused on

- identifying and prioritizing the most impactful networking improvements to reduce carbon emissions;
- defining action items for a carbon-aware networking research agenda; and
- fostering and facilitating research collaboration in order to reduce carbon emissions and to positively impact climate change.

As a major result from the seminar, the following problems and topics for future research have been identified:

1. characterizing the Internet footprint on carbon emissions accurately;
2. understanding attributional and consequential accounting of carbon emissions in networked systems; and
3. identifying potential solutions to give network systems more flexibility in better supporting energy grids and connecting to renewable energy sources.

One of the concrete results of this seminar is a list of technologies and research opportunities for which we estimated the potential impact and time horizon.

Overview of the Seminar: Motivation and Conceptual Framing

What is the path towards a Net Zero Internet? Reducing CO₂ emissions to combat climate change is one of the greatest challenges facing mankind. An overarching goal is to make human activities sustainable, where sustainability is defined as: *Meeting the needs of the present without compromising the ability of future generations to meet their own needs (UN Brundtland Commission, 1987).*

Networking technologies are a key part of this challenge – as a provider for solutions (such as enabling teleworking and reducing travel), but also as a contributor to CO₂ emissions, for example, through significant power consumption. Thus, it is becoming important to make networks themselves “greener” and devise solutions that result in less carbon intensity while continuing to meet the increasing network demands and service requirements.¹ Complicating the matter, energy consumption is related to CO₂ emissions in ways that vary greatly by network operators (through the use of “clean” energy).

Many of today’s improvements are driven by general advances in computing hardware as well as in transmission technology (antennas, lasers). While this is where the biggest opportunities for energy reduction may lie, it is important to extend questions of greenness

¹ Reducing traffic demands could potentially reduce network emissions. However, this is a social issue that was considered outside of the scope of this seminar.

to other aspects of carbon reduction, both temporal and spatial, and to other layers in the networking stack – to the data and control plane, to routing and traffic forwarding, to the ways in which networks are organized and deployed. For example, can network protocols be designed in ways that make them more carbon-efficient, e.g. by routing along a path that is primarily powered by renewable energy, so as to reduce CO₂ emissions with proper path selection? Can control planes be designed in a way that allows to rapidly switch off resources that may currently not be needed without compromising other important goals such as network resilience or security? How much overhead is introduced by current architectures and protocols for cloud computing, media streaming, and content distribution, and what is the energy conservation potential of alternative systems? What protocol advances could enable greener networking solutions? How can networks be optimized not just for QoS or utilization but for low environmental impact?

This Dagstuhl Seminar brought together researchers of different perspectives with a vested interest on this topic, to exchange ideas and explore the possibilities for synergy in their respective research directions. We identified and prioritized the most impactful networking improvements to reduce carbon emissions, defined action items for a carbon-aware networking research agenda, and fostered/facilitated research collaborations that result in improved techniques to reduce carbon emissions.

This seminar was organized into sessions pertaining to these themes, with presentations from leading researchers drawn from the participants. Researchers in the six main topics – devices, protocols, metrics, measurement, architecture, and management – were invited to participate.

The seminar and this report follow the same structure: Section 3.3 discusses the bigger picture of energy saving and carbon neutrality for applications, systems, and stakeholders. Section 4.4 discusses potential technical energy-saving approaches for networking functions such as routing, traffic engineering, power saving mechanisms, and technology-specific mechanisms, such as for wireless communications. Section 5.4 discusses overarching concepts in technology related to cyclical characteristics, such as control loops, i.e., the application of control logic to optimize environmental sustainability of solutions, as well as the actual life-cycle of network equipment. Finally, Section 7 summarizes the outcomes of the seminar in three tables listing different proposed research directions.

We acknowledge that other topics and perspectives could have been discussed, and that this selection is a result of many parameters: the expertise of the participants, the subjectivity of the organizers, or the amount of time available for discussions. Other ways to frame the conversation could be the basis for another future seminar.

2 Table of Contents

Executive Summary

Alexander Clemm, Dirk Kutscher, Michael Welzl, Cedric Westphal, and Noa Zilberman 168

Talks

| | |
|---|-----|
| Mobile Networking Perspective <i>Jari Arkko</i> | 171 |
| Greening of the Internet – The Big Picture <i>Romain Jacob</i> | 172 |
| Application, Systems, Stakeholders | 172 |

Talks

| | |
|---|-----|
| Sustainable ICT <i>Daniel Schien</i> | 177 |
| Streaming & Carbon emission <i>Te-Yuan Huang</i> | 177 |
| Green Video Streaming <i>Farzad Tashtarian</i> | 178 |
| Network Technologies | 178 |

Talks

| | |
|---|-----|
| On the Duality of Power System and ICT <i>Hermann de Meer</i> | 181 |
| Carbon-Aware Routing: Metrics and Challenges <i>Sawsan El Zahr</i> | 182 |
| Greening the Internet with SCION <i>Jonghoon Kwon</i> | 182 |
| Lifecycle and Control Loops | 182 |

Talks

| | |
|---|-----|
| Energy aware orchestration and lifecycle management of Telco workloads: an NFV-MANO approach <i>Hammad Zafar</i> | 186 |
| Pathways for sustainability on transport networks <i>Luis Miguel Contreras Murillo</i> | 186 |
| For a joint operator-centric approach to assessing network management effort <i>Marc Bruyère</i> | 187 |
| Autopower and Network Power Zoo <i>Romain Jacob</i> | 187 |

| | |
|----------------------------|-----|
| Seminar outcomes | 187 |
|----------------------------|-----|

| | |
|------------------------|-----|
| Participants | 192 |
|------------------------|-----|

3 Talks

3.1 Mobile Networking Perspective

Jari Arkko (*Ericsson – Jorvas, FI, jari@arkko.com*)

License © Creative Commons BY 4.0 International license
© Jari Arkko

This talk discussed the current position regarding the energy consumption of mobile networks; what parts of the network are the most power hungry; and what avenues exist for improvements for mobile networks.

Studies of network operators all over the world indicate that overall, the power consumption of networks is on a modest increase. These studies also show that there is little correlation between the amount of data transported by the networks and the amount of energy they consume. It is remarkable considering that the amount of data has increased by several orders of magnitude in the last twenty years. Nevertheless, improving the power consumption of the networks is a key goal from an overall sustainability point of view, from the goals of reducing ICT power consumption in general, but also because energy is a significant cost for operators. There is considerable demand for improvements in this area.

The radio network consumes the most energy in mobile networks, largely due to the use of power amplifiers, the required cooling elements, etc. Given such a dominant factor, this has been the focus of sustainability improvements. Mobile networks are also characterized by usage patterns that fluctuate widely in terms of time and location. Since the mobile network is fundamentally a function that is tied to a location, this also leads to relatively low average load rates. Consequently, achieving energy proportionality² is an important missing element.

One particularly beneficial area of improvements relates to the ability of networks to be in a sleep mode for a greater fraction of the time, given the amount of traffic fluctuates. The sleep mode could dynamically impact a small subset of nodes for a short period of time, or some entire cells could be turned off. The former is possible because of the nature of the mobile networks as highly controlled and scheduled systems. The latter is possible due to overlapping spectrum areas and cells, making it possible to turn off, for instance, small cells without impacting the ability of anyone to get connectivity from the remaining macro cells.

The ability of the systems to benefit from the brief periods of sleep depends on two factors. One, the radio standards need to be designed so that they do not impose an excessive requirement for always-on-operation, e.g., due to too frequent synchronization, reference, and system information signals. And two, the ability of implementations to leverage opportunities created by standards to actually turn themselves off, for instance, by being able to control over short timescales whether the power amplifiers are on. Improvements from LTE to 5G have already made significant progress in these two aspects, and further progress seems feasible in the coming network generations.

² Energy proportionality is achieved when the energy consumed by the network is proportional to the utilization.

3.2 Greening of the Internet – The Big Picture

Romain Jacob (ETH – Zurich, CH, jacobr@ethz.ch)

License  Creative Commons BY 4.0 International license
© Romain Jacob

This talk provided a general overview of the sustainable networking context. A key point of the presentation was to emphasize the importance of clarifying the scope of the problem or technical solution we have in mind when discussing “the Internet”. Some important dimensions to be aware of include:

- the particular portions of the overall network: which part of the network can, or needs to, be optimized, the core network, the access network, the metro area network, the datacenter.
- the network layer: from L1 to L7
- the class of protocols, inter-domain vs intra-domain
- the target users: hyperscalers, ISPs, enterprises, or end-users
- the agent of change: technological, behavioral or regulatory.
- the progress strategy: better proportionality, raw improvement, waste reduction, resource use optimization, et.
- the metrics: energy usage or carbon emissions, in an relative or absolute measurement
- the life cycle phase: operational or including manufacturing and recycling as well (that is, taking into account the embodied carbon)
- the class of reasoning: attributional or consequential

Attributional represents the emissions that can be attributed to one actor in particular, for instance by dividing the total amount of emissions by the number of participants. Consequential represent the activities that can be associated as a consequence for an activity. Consequential reasoning weighs the pros and cons of decisions often in terms of total carbon emissions rather than focusing on how to allocate the responsibility to each party. This taxonomy can be used to facilitate the conversation in order to define the proper scope.

3.3 Application, Systems, Stakeholders

Networks exist within larger systems, such as large-scale computing infrastructure and user-facing services. Effective research agendas for energy savings and carbon neutrality should hence take a broader perspective and understand the bigger picture, including identifying the most energy-intensive sub-systems. In addition, the network and the larger networked systems interact with the energy grid and renewable energy sources in different ways. This raises questions on how these systems and potential energy saving measures interact with respect to the requirements that they impose on each other.

The discussions at the seminar included energy footprint, potential architectural issues, and research challenges in relevant larger systems such as the Web, Media Streaming, Mobile Communications, and AI workloads. We discussed specific systems, problems, requirements, problems, solution approaches, associated timelines, and assessment of potential impact.

In this section, we first talk about conceptualizing environmental effects of ICT (section 3.3.1) and about different categories of technical efforts (section 3.3.2). We then discuss specific systems such as video streaming (section 3.3.3), mobile communication networks (section 3.3.3), and AI workloads (section 3.3.5).

3.3.1 Conceptualizing Environmental Effects of ICT

According to a recent EU report on *Energy Consumption in Data Centres and Broadband Communication Networks in the EU* [1], EU data center energy consumption is slowly increasing.

According to a 2020 GSMA report titled *COVID-19 Network Traffic Surge Isn't Impacting Environment Confirm Telecom Operators*³, the energy proportionality of the Internet is relatively low. For example, UK operator BT reported a 100% increase in daytime traffic across its fixed broadband network during COVID-19 times. However, this did not lead to a noticeable increase in electricity use or carbon emissions.

Reference [2] reports on *the overlooked environmental footprint of increasing Internet use* and states that, “depending on the energy supply mix and use efficiency, Internet traffic contributes differently to negative environmental impacts and climate change”. The report also claims that “as the number of Internet users increases, the number of online services and applications they use grow” [2]. This trend exacerbated the environmental footprint of the Internet, which can be observed through increased usage of Internet services such as video streaming (see section 3.3.3).

The accounting of carbon emission should distinguish between *attributional* and *consequential* emissions⁴. Attributional accounting estimates the share of the total impacts of a system that can be attributed to a specific function. For instance: what is the share of the energy spent in the Internet for video streaming. Consequential accounting on the other hand considers the environmental impact induced by the implementation of this function. It compares a world without this function (say, without video streaming, where you drive your car to the movie theater) with a world with this function (say, with video streaming, where you watch a movie from home).

Traditional attributional accounting is often applied in reporting the carbon emission impact of an isolated unit, such as a company or specific sector. However, energy saving and carbon reduction measures in one unit or sector could lead to an overall increase in the larger socio-economic system, for example when a data center switches to renewable energy, while the overall amount of available renewable energy does not change, so that other consumption, with greater flexibility potential, would then use less carbon-neutral energy.

Details can be found in the presentation summarized in section 4.1.

3.3.2 Different Categories of Technical Efforts

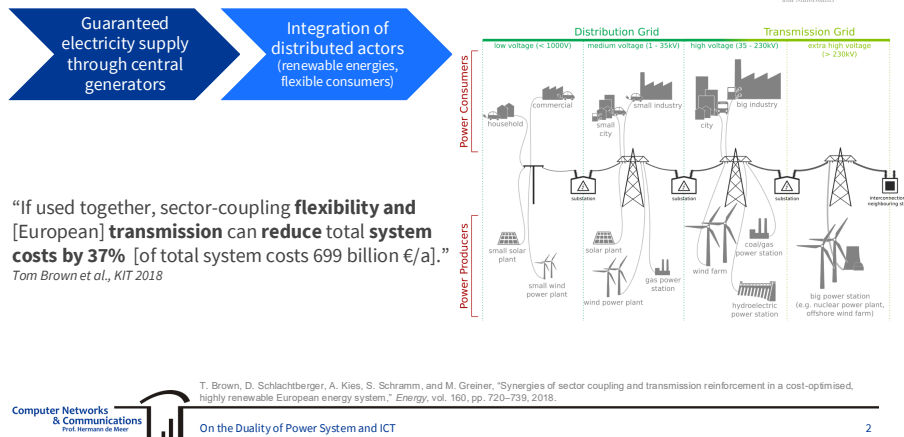
There are different *potential energy efficiency improvements*, such as more efficient link layer communication, load aggregation and system consolidation, utilization-proportional energy consumption, load-adaptive system designs, and shifting workloads in ways that optimize energy efficiency, for example by concentrating workloads on fewer systems.

These general energy efficiency improvements can generally lower the overall energy consumption over time, for example per year, and they can *increase energy flexibility potentials* by offering higher levels of adaptability. However, increased levels of adaptability and myopic optimization can also lead to *problematic flexibility requirements* imposed on the energy infrastructure, for example when workloads and energy needs are concentrated, leading to unwanted spikes under a certain energy provisioning contract.

³ <https://tinyurl.com/4kva7dmb>

⁴ <https://hubblo.org/blog/attributional-vs-consequential/>

Motivation: From „Supply follows Demand“ to a flexible energy system



■ **Figure 1** Sector-Coupling Energy Reduction Potential.

For an effective holistic optimization, it would be important to match energy requirements and saving potentials, for example through a mediator. As reported by Brown et al. [3] and as illustrated in Figure 1, it is estimated that, in Europe, sector-coupling flexibility and transmission can reduce total system costs by 37%.

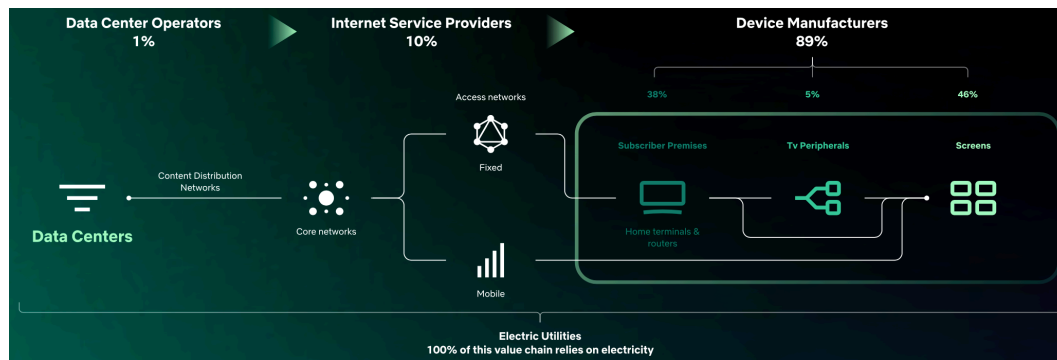
In general, predictability is considered to be very valuable for the power grid. Optimizations that are not expected by the power grid and that are unaware of it can have a negative impact. Such local optimizations can create an undesirable need for increased flexibility from the power grid. It is deemed better to have sensitive optimization informed by a holistic system view.

In the talk described in section 5.1, an example was given, where an optimization significantly reduced energy, but as a result, the power price increased significantly. This is because temporary power surges are considered to have a particularly bad impact on the grid; in the short term, avoiding them, and “flattening” the use may sometimes be more valuable than reducing the average energy usage. Nevertheless, in the long term, it is an important goal to reduce the average.

3.3.3 Video Streaming

Te-Yuan Huang (Netflix), also see section 4, presented facts on *Carbon Emission and Streaming* at Netflix. According to measurements at Netflix, 89% of the energy in the streaming pipeline is consumed by user devices (TVs and mobile viewing platforms, as well as home routers and peripherals), 10% is consumed by the core and access network, and only 1% by data centers. However the whole streaming pipeline only consumes about 3% to 5%, according to a recent Netflix Environmental Social Governance Report [4]; a breakdown of what these 3-5% consist of is shown in Figure 2.

In more absolute terms, in the UK, the carbon footprint of one hour of video streaming equates approximately 55 gCO₂e (grams of carbon dioxide equivalents), which is roughly equivalent to three boils in an electric kettle [5, 6]. Details can be found in the presentation summarized in section 4.2.



■ **Figure 2** Netflix Streaming Pipeline Energy Consumption [4].

Farzad Tashtarian from the Department of Information Technology, Alpen-Adria-Universität, Klagenfurt, Austria, presented Coconut [7], a COntent COnsumption eNergy measUrement daTaset for adaptive video streaming collected through a digital multimeter on various types of client devices, such as laptop and smartphone, streaming MPEG-DASH segments. The research in the paper presents perspectives on the energy usage of streaming devices⁵.

With respect to assessing the power consumption of the streaming pipeline, the following research questions were proposed (see section 4.3):

- How can we accurately measure the energy consumption of video streaming components in real-world scenarios?
- To what extent can Generative AI impact the energy consumption of a video streaming pipeline?
- What is the effect of emerging protocols (e.g., QUIC) on energy consumption?
- How can we address the energy consumption challenges of immersive media?
- How can we balance latency, QoE, and energy efficiency effectively?
- What strategies can be employed to characterize and potentially improve the energy-efficiency of Adaptive Bitrate (ABR) algorithms?
- Can caching strategies be used to improve energy-efficiency?

3.3.4 Mobile Networks

Most of the energy optimization solutions on the mobile network side are dealing with temporarily shutting down nodes/sites or moving major workloads across sites and segments of the network on the system level. This has a strongly interlinked relationship with grid dynamics on the power system side.

Information exchange and a feedback loop between the two systems (mobile network and power grid) can be helpful in studying the impact of optimizations (on the network side) on power system dynamics on the grid. In some cases (small network sites), the impact could be minimal (hopefully) but in case of planning large workload deployments across the system, the impact can be significant. After understanding the potential impact of network side optimizations on the grid, further autonomous, fine-grained optimizations on

⁵ Dataset available at <https://athena.itec.aau.at/coconut/>

the site/cluster/data center level can be made to make the cloud infrastructures and orchestration/scheduling systems more energy efficient. Details can be found in the presentation summarized in section 3.1.

3.3.5 AI Workloads

There is significant potential to make AI, both training and inference, more efficient in itself. The seminar also discussed the need of making AI systems aware of and responsive to larger system constraints.

Much of the focus in networking for AI workloads has been on making machine learning faster, without much consideration for energy efficiency, which is in part explained by the lack of data to assess emission impact of different training models. Hence, providing more comprehensive instrumentation of networking devices and data center equipment will be an important enabler to make AI workload placement more energy-efficient.

AI workloads generally also involve streaming of various types of telemetry data that is used to train and apply AI models from the sources of that data to where it can be processed. Some of the lessons from the greening of video streaming thus potentially also apply to the greening of the handling of AI data, both input and output, for inference usage, despite some obvious differences such as the number of clients to stream to (few in the case of AI processing, many in the case of video) or the direction of streaming (potentially from the edge to the core unlike from the core to the edge as with video). However, it may make sense to apply concepts from video such as DASH (Dynamic Adaptive Streaming over HTTP) and application of different video codecs (to send images at different resolutions, frame rates, and color depth depending on current network conditions) also to the streaming of AI data.

One approach to minimize energy consumption by AI workloads consists of moving workloads close to the sources of energy. One example is a project in Germany that aims to host compute nodes and mini-data centers in windmills. As 1/3 of energy is lost in the energy grid, moving the workload to the source will make AI workloads correspondingly more energy-efficient. This requires the ability to adapt workloads to available energy and renewable energy source dynamics, as well as applications that are flexible in terms of their workload processing demands. Similarly, the concepts of edge intelligence and federated learning can play an important role in allowing AI workloads that can be run across a network more energy efficient. Traditionally applied to minimize latency (by processing data close to the source) and improve privacy (by not requiring data to leave the edge), this provides for a promising additional incentive to be considered.

References

- 1 G. Kamiya and P. Bertoldi. Energy consumption in data centres and broadband communication networks in the eu. Technical report, European Union, 2024.
- 2 Renee Obringer, Benjamin Rachunok, Debora Maia-Silva, Maryam Arbabzadeh, Roshanak Nateghi, and Kaveh Madani. The overlooked environmental footprint of increasing internet use. *Resources, Conservation and Recycling*, 167:105389, 2021.
- 3 T. Brown, D. Schlachtberger, A. Kies, S. Schramm, and M. Greiner. Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable european energy system. *Energy*, 160:720–739, 2018.
- 4 Netflix. Environmental social governance report. Available at https://s22.q4cdn.com/959853165/files/doc_downloads/2024/6/2023-Netflix-Environmental-Social-Governance-Report.pdf, 2023.

- 5 The Carbon Trust. Carbon impact of video streaming. *Available at* <https://www.carbontrust.com/our-work-and-impact/guides-reports-and-tools/carbon-impact-of-video-streaming>, June 2021.
- 6 Daniel Schien, Paul J S Shabajee, James F Wickenden, William Picket, Glynn Roberts, and Chris W Preist. Demo: The dimapct tool for environmental assessment of digital services. In Engineer Bainomugisha and Waylon Brunette, editors, *Proceedings of the 4th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies, COMPASS 2022*, ACM International Conference Proceeding Series, pages 701–703, United States, June 2022. Association for Computing Machinery (ACM). Publisher Copyright: © 2022 Owner/Author.
- 7 Farzad Tashtarian, Daniele Lorenzi, Hadi Amirpour, Samira Afzal, and Christian Timmerer. Coconut: Content consumption energy measurement dataset for adaptive video streaming. In *Proceedings of the 15th ACM Multimedia Systems Conference, MMSys '24*, page 346–352, New York, NY, USA, 2024. Association for Computing Machinery.

4 Talks

4.1 Sustainable ICT

Dan Schien (University of Bristol, GB, Daniel.Schien@bristol.ac.uk)

License © Creative Commons BY 4.0 International license
© Daniel Schien

This talk provided an overview of the environmental effects of information and communication technology (ICT) through a lifecycle perspective. After briefly distinguishing the direct and indirect contributions of the ICT sector to global greenhouse gas emissions, including electricity consumption in data centers and networks, it looked at the relative scale of the contribution of the ICT sector to global greenhouse gas emissions, before focusing specifically on the contribution by networks.

The talk highlighted that the relationship between the use of networks and the energy consumption is complex. It also considered the complex interactions with the carbon emissions from the electricity consumption and provided a categorisation of existing assessment methodologies for product carbon footprinting of digital services.

4.2 Streaming & Carbon emission


Te-Yuan Huang (Netflix – Los Gatos, US, thuang@netflix.com)

License © Creative Commons BY 4.0 International license
© Te-Yuan Huang

This talk examined carbon emissions in various stages of the streaming pipeline, using Netflix as an example case study. The data of Netflix shows that the Network accounts for approximately 10% of the user-phase streaming carbon emission. It's a small but still sizable portion. Netflix noticed that static power consumption on routers and servers accounts for a big portion of the network's emissions, and energy consumption does not grow with data traffic on fixed networks. There are still a lot of unknowns in this field—how do we do more measurement studies for mobile networks?; how do we reduce static power consumptions?; and many more. This talk presented what Netflix has learned and the questions the company would love to answer together with the community.

4.3 Green Video Streaming

Farzad Tashtarian (Alpen-Adria-Universität – Klagenfurt, AT, farzad.tashtarian@aau.at)

License  Creative Commons BY 4.0 International license
© Farzad Tashtarian

This talk presented the work at the Christian Doppler Laboratory ATHENA on COCONUT, a testbed for fine-grained, real-world energy measurements across video streaming stages. It covered an analysis of the energy impacts of frame rate, luminance, and other parameters, exploring strategies to optimize energy use while balancing quality and sustainability. The insights aim to enable eco-efficient streaming without compromising user experience.

4.4 Network Technologies

The previous sections have discussed the systems aspects of green networking. This section focuses on specific network technologies, including functions such as routing, traffic engineering, power saving mechanisms and technology specific mechanisms, such as for wireless communications.

For each function or technology, several questions have to be asked, including:

- What is the technology's energy and carbon footprint?
- What is the saving potential?
- What are the requirements to enable the green technology?
- What are the challenges to do so?
- What is the timeline and what will be the impact?

Some of these questions were later readdressed in the wider context of the seminar.

In this seminar, the current state in power saving mechanisms was introduced by Noa Zilberman. Wide area network traffic tends to have periodic trends, with peak traffic during evening time and very little traffic during night time. However, the power consumption of the network is not proportional to these trends.

Power proportionality is a long standing problem in computing and networks [1]. While ideally a device's power consumption will directly correspond to its utilization, in practice 70%-95% of routers' energy consumption is static, independent of load. In contrast, CPUs have implemented mechanisms that allow them to be more (but not entirely) power proportional. The reason for this difference is the strict requirement from network devices not to drop traffic, whenever it arrives. Industry has tried for years to improve network devices' power efficiency, achieving 90% improved power efficiency per 100G, but at the same time bandwidth has increased by a factor 80 [2]. Even on an entire ISP level there has been 8.6% energy reduction despite x8.6 increase in traffic [3]. Wired-focused power saving mechanisms in production range from architecture (higher radix devices) and turning off links or cards to adjusting link rate. On the device level, we see higher integration, clock and power gating, and power throttling as saving mechanisms. These all come to show that power reduction mechanisms are already implemented in network devices, and have shown significant savings. However, these savings are not enough given the overall traffic increase, and we need innovative ideas that result in significant improvements while also conforming with operational network needs.

Four breakout groups have been formed to discuss different technologies: Power saving mechanisms, routing, traffic engineering and congestion control, and wireless.

4.4.1 Power Saving

This discussion started from the joint optimization of the power grid and ICT, as increasing ICT energy efficiency can affect the grid's proper operation. This means that ICT needs to consider power grid constraints and multi-disciplinary efforts in joint optimization may be required to achieve meaningful improvements. A second key take-away is that visibility about demand cycle, peak times and pricing models would be necessary for data center operators to consider. This would allow to match traffic workload forecasts with energy forecasts. On the other hand, it may be possible to use power in off-peak hours to scavenge work-loads in DC.

An additional discussion considers the optimal use of resources. It starts with scheduling in the operating system and orchestration layer, and extends to the scheduling of resources from an existing resource pool. For example, is it better to use one CPU at high utilization, or to use more CPUs in lower utilization? One observation from telecommunication operators is that CPU offloading is happening more and more, with GPUs and smartNICs in use. How to make these offload targets more energy efficient or use software replacement remains a challenge. It is also clear that the right hardware needs to be identified for different tasks, such as the optimal selection of access network technology. For instance, passive optical network (PON) can reduce energy in transport, but at the same time using optics can result in energy waste if used for short distance.

The discussions emphasized concerns that reduction in power usage and improvements in energy efficiency during operations do not necessarily imply greater sustainability. Instead, to optimize greenness, the parameter to minimize concerns the maximum power drawn at any one time. The reason is that the supply of power must be dimensioned to be able to satisfy that maximum power draw. It is this parameter, which determines the amount of electricity that needs to be supplied and hence generated, in turn correlated with carbon generation due to the inclusion of fossil fuels in the power mix. Reduction in power usage (on the scale of MWh) is thus beneficial only when planned sufficiently in advance for the grid provider to adjust to it. From the grid provider's perspective, power that is not being put to productive use is effectively wasted as it still generated. In some cases, improvements in energy efficiency may even be counterproductive if reduction in energy use during some periods is compensated by additional spikes during peak periods.

One of the points brought up during the discussion concerns the role of prices as a mechanism to steer behavior of customers. Technical measures to improve environmental sustainability should be accompanied with price schemes to provide the right incentive structure to steer solutions in a way that truly reduces the amount of energy that is not just used but that needs to be supplied.

Also touched upon was the aspect of energy ratings of devices. Schemes such as Energy Star (used to allow for informed purchase decisions regarding the energy efficiency of electrical appliances) are not supported for networking equipment. While there are some questions regarding the efficacy of such schemes, ready availability of such data appears beneficial. To be meaningful, this would need to be provided in an objective manner. Today, it is not uncommon for data that is being advertised by equipment suppliers, or that is provided by equipment's own power measurement instrumentation, to deviate significantly from data provided by objective measurements.

4.4.2 Routing

One of the areas already being explored is carbon-aware routing. However, insight from SCION's work[4] will be hard to integrate with BGP. It is also clear that security should not be compromised, and it was asked if SCION multi-path will create more points of failure.

Using energy labels [5, 6] was noted as an idea for sharing information that might work across domains. It was further noted that peering agreements can be used to improve the transparency of green path selection procedures. One of the questions is the hot potato routing problem. This is not a new problem, but the local optimization of carbon efficiency is an incentive for companies to pass packets as soon as possible to other ASes. Another aspect to consider is the actual distribution of traffic types in the backbone and potentially the need to differentiate traffic. From industry partners contributions, there appears to be significant differences between countries and networks, such as the existence or lack of video streaming caches within a network, which may lead to increased load. A takeaway is that sharing information across ASes is a big challenge, requiring validation that is not necessarily network based and potentially a third party to regulate or monitor. It was proposed that federations might be able to enforce it. Deciding on metrics that can be shared across domains remains an open research question.

4.4.3 Traffic Engineering (TE) and Congestion Control (CC)

Carbon aware routing and carbon aware TE potentially changes the location and severity of network congestion. Lower than Best Effort (LBE) is a type of traffic that can be less latency sensitive [7, 8]. LBE can be used for elastic traffic that can be treated more flexibly in the network (unlike streaming), e.g. to send it over a longer but more carbon efficient path. Recent work in CC considered pacing traffic, spreading it over time [9]. This conflicts with desirable wireless operation, where the objective is to send more traffic in bursts, in order to sleep more. There are some other conflicts, such as buffering, which can also introduce waste. There are important trade-offs: a need to provision, but also to avoid waste. A significant challenge in traffic engineering is how to benefit from renewable energy without exhausting its resources and requiring non-renewable energy to be used. There must be a control loop to guarantee this. On the other hand, previous work indicated that the power consumption of networking equipment is below the threshold to trigger such energy-source transition. A potential future CC challenge is AI traffic, which was already shown to be a bottleneck for frontier AI models training within data centers. However, it is unknown what the effect in wide area networks will be. Clearly, inference traffic differs from training and may look a lot like other types of network traffic. Furthermore, the AI tasks will affect the nature of the traffic: from object recognition from camera feeds to ChatGPT text prompts. It is also possible that new applications will require the use of training AI models over the wider network. Privacy-preserving A.I. training and inference schemes like federated learning and split learning may move load further out of the data centers into the WAN, necessitating resource allocation schemes sensitive to the energy usage patterns they affect.

4.4.4 Wireless and mobile communication

The granularity of mobile and wireless communication is different from wired. For example, it is possible to turn off base stations at night or in certain city areas. It is also possible to turn off components within each base station, such as power amplifiers and antennas. However, the impact on transport and application layer would require further studies. There is a tension between what is ideal in wired, and what is ideal in wireless (such as the previously mentioned pacing). Burstiness is good in wireless, but not in wired. Wireless networks may look at demand shifting and using the time dimension to cache contents; or shift mobile traffic onto wifi (offloading). A question was raised about the introduction of 6G, using smaller cells, higher frequencies. Would it be better to turn off smaller cells and use larger

ones to cover with better efficiency? One power saving mechanism is to turn off equipment. In mobile, this ranged from the whole base station to only switching off the power amplifiers. Shutting down the whole base station may lead to aging concerns of the equipment with cooling down and powering up. Another important concern is management: once you turn something off, you have no visibility to it. Therefore, the management layer is afraid of not just losing visibility, but that once off, the resource won't turn back on when required.

References

- 1 Dennis Abts, Michael R Marty, Philip M Wells, Peter Klausler, and Hong Liu. Energy proportional datacenter networks. In *Proceedings of the 37th annual international symposium on Computer architecture*, pages 338–347, 2010.
- 2 Itzik Kiselevsky. Evolution of switches power consumption. Presented at the Carbon Aware Networks Workshop 2023, September 2023.
- 3 Telefonica. ESG Highlights, Annual Report 2023. Technical report, 2024.
- 4 Seyedali Tabaeiaghdaei, Simon Scherrer, Jonghoon Kwon, and Adrian Perrig. Carbon-aware global routing in path-aware networks. In *Proceedings of the 14th ACM International Conference on Future Energy Systems*, pages 144–158, 2023.
- 5 Noa Zilberman, Eve M Schooler, Uri Cummings, Rajit Manohar, Dawn Nafus, Robert Soulé, and Rick Taylor. Toward carbon-aware networking. *ACM SIGENERGY Energy Informatics Review*, 3(3):15–20, 2023.
- 6 Sawsan El-Zahr, Paul Gunning, and Noa Zilberman. Exploring the benefits of carbon-aware routing. *Proceedings of the ACM on Networking*, 1(CoNEXT3):1–24, 2023.
- 7 David Ros and Michael Welzl. Less-than-best-effort service: A survey of end-to-end approaches. *IEEE Communications Surveys & Tutorials*, 15(2):898–908, 2013.
- 8 Stanislav Shalunov, Greg Hazel, Jana Iyengar, and Mirja Kühlewind. Low Extra Delay Background Transport (LEDBAT). RFC 6817, December 2012.
- 9 Michael Welzl, Wesley Eddy, Vidhi Goel, and Michael Tüxen. Pacing in Transport Protocols. Internet-Draft draft-welzl-icrg-pacing-01, Internet Engineering Task Force, October 2024. Work in Progress.

5 Talks

5.1 On the Duality of Power System and ICT

Hermann de Meer (University of Passau, DE, hermann.demeer@uni-passau.de)

License © Creative Commons BY 4.0 International license
© Hermann de Meer

This talk advocated for an integrated view and systemic approach of ICT load optimization and power system flexibility potentials. Taking power system flexibility potentials explicitly into account avoids possible pitfalls of isolated optimization decisions that may have detrimental effects from the perspective of sustainability. Under certain assumptions, isolated ICT optimization decisions might neither lead to desired power savings nor to carbon food print reductions. On the contrary, from the systemic point of view, detrimental effects may result: a counterintuitive increase of power demand or/and an increased carbon food print. A systematic approach based on sector coupling could mitigate these detrimental effects.

5.2 Carbon-Aware Routing: Metrics and Challenges

Sawsan El Zahr (University of Oxford, GB, sawsan.elzahr@eng.ox.ac.uk)

License  Creative Commons BY 4.0 International license
© Sawsan El Zahr

Every click, swipe and scroll leaves an impact on the environment. As we rely more on the Internet in many sectors and for many applications, the carbon footprint of the internet is rising, and it is unclear if we can achieve the net zero goals by 2050. Introducing carbon awareness to computer networks is one promising solution, yet with many challenges. The work presented in this talk is in the context of fixed wired networks, where accounting for their emissions is hard, requires changes to deployed equipment, and has contentious benefits. The talk has shed light on the benefits of carbon aware networks, by exploring a set of potential carbon-related metrics and their use to define link-cost in carbon-aware link-state routing algorithms. Using realistic network topologies, traffic patterns and grid carbon intensity, it identified useful metrics and limitations to carbon emissions reduction. Consequently, a new heuristic carbon-aware traffic engineering algorithm, CATE, was proposed. CATE takes advantage of carbon intensity and routers' dynamic power consumption, combined with ports power down, to minimize carbon emissions. The presented results show that there is no silver bullet to significant carbon reductions, yet there are promising directions without changes to existing routers' hardware. The talk uncovered some of the challenges that Internet Service Providers (ISPs) will need to face as we move towards net-zero networks.

5.3 Greening the Internet with SCION

Jonghoon Kwon (ETH – Zurich, CH, jonghoonkwon@gmail.com)

License  Creative Commons BY 4.0 International license
© Jonghoon Kwon

The increasing demand for energy efficiency in network operations calls for innovative solutions to reduce carbon emissions while maintaining performance. This talk explored how SCION, a next-generation Internet architecture, can facilitate greener Internet practices through its path-aware and multipath routing capabilities. By integrating energy and CO₂ metrics into path selection, SCION enables energy-optimized routing and empowers applications to choose routes based on environmental impact. This presentation highlighted ongoing research opportunities in integrating energy-efficient strategies with SCION and invited collaboration through the SCION Education, Research, and Academic Network (SCI ERA).

5.4 Lifecycle and Control Loops

5.4.1 Overview

This section focuses on two separate yet related aspects of greening networking, sharing the fact that many overarching concepts in technology resemble cyclical characteristics when viewed from a holistic perspective. One of these aspects concerns control loops, i.e. the application of control logic to optimize environmental sustainability of solutions. The other aspect concerns the lifecycle of network equipment, which encompasses not only its operation

but a larger cycle that also includes aspects such as manufacturing and decommissioning of equipment, all of which contribute to the environmental footprint and thus deserve consideration.

Three breakout groups were formed to discuss different aspects of these aspects: **1. Observations**, **2. Planning and Forecasting**, and **3. Hardware Lifecycle**. The following sections present the outcomes of these conversations.

5.4.2 Control Loops

Control loops are a universal principle that plays a role in many solutions that aim to improve outcomes by continuously observing (the entity or domain that is being controlled and parameters that are being optimized), analyzing (to interpret and make sense of the observations, perhaps correlating them with other data), deciding (on which course of action to take in search of an improvement), and acting (i.e. performing the steps of the course of action).

Control loops also apply to many potential solutions that aim to improve green networking outcomes – for example, aiming to minimize the amount of energy used to provide communication services while meeting other important operational goals such as adhering to service level objectives or ensuring network resilience.

In the context of green networking, the following are important considerations to provide control loop components that will facilitate better solutions:

Observation

- Which observations will facilitate solutions? Specifically, which observations specifically related to carbon, energy usage will be needed?
- Which metrics will define success?
- Which instrumentation (of the network, of devices) will therefore be needed to obtain those metrics?
- How will measurements be performed?
- How will observations be collected?
- How can it be ensured that observations are in fact accurate, that they can be trusted when relied upon in solutions, that they are timely?

One particular topic deserving further consideration concerns observations in virtualized environments, specifically how to fairly attribute real-time power measurements from the underlying hardware layer to virtualized entities such as a Virtualized Networking Functions (VNFs).

Another important aspect concerns the need to exchange carbon metrics across boundaries of organizations. Visibility into data across domains facilitates holistic solutions, however, network providers are reluctant to share such data due to privacy and security implications. Absent effective mechanisms and trust relationships to address those issues, solutions will be confined to work within single domains.

Among metrics, one promising avenue of exploration concerns support for path metrics that allow to assess energy consumption and carbon intensity end-to-end [1]. Such metrics can in turn be used to make informed carbon-aware traffic steering decisions and to lead to carbon-optimization of configured paths.

Analysis

- How does data need to be aggregated in ways that are meaningful to a solution? What is the “big picture” conveyed by the data?
- What conditions indicated by data are “significant”, to facilitate prioritization of areas of focus?
- Can forecasting for certain key parameters be improved? The ability to make better forecasts (for anything from traffic demand to energy usage) may improve the effectiveness of solutions.

Decision

- What algorithms will reduce energy usage and carbon footprint?
- How can general techniques (such as Generative AI or Intent-Based Networking) be applied to green use cases? What customizations, what custom algorithms need to be developed?
- What tools, what decision support systems will help network operators (and vendors, and end users) make better green decisions?

Action

- Do we have the right set of control knobs that can be used to affect carbon footprint in a meaningful way? (For example, this may include support for meaningful power saving modes or downspeeding.)
- Which additional mechanisms might be needed to facilitate such control knobs? (For example, this may include mechanisms to facilitate discovery and state convergence as resources are dynamically made available.)

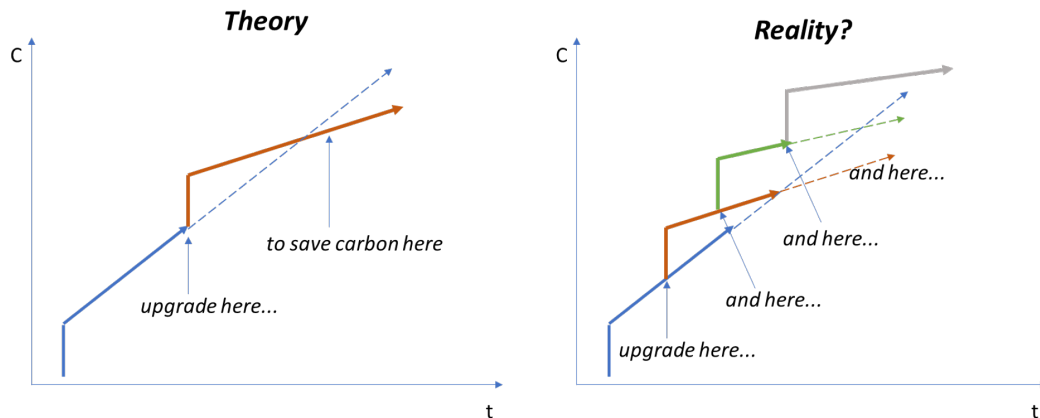
5.4.3 Lifecycle

The natural starting point for greenness considerations involves the operation of a network, which involves the need to power equipment or to provide cooling and ventilation. However, greening the network does not start there – before equipment can be taken into operation, it needs to be manufactured. After equipment is decommissioned, it needs to be disposed of. In other words, there is a lifecycle for equipment that is to be considered. A substantial part of the green equation concerns the carbon that is embodied in a device itself, including energy use for manufacturing, not even to speak of use of land and water used to extract raw materials.

As a result, networks are greener if they can avoid deployment of hardware that is not really required. Today, networks are underloaded in many cases, meaning that the same service levels could arguably be achieved with less buildout [2, 3]. One reason for this is that overprovisioning is commonly used as a way to avert potential problems. Overprovisioning may be cost effective, but not necessarily “green”. Hence, a key question concerns when network upgrades are really required. By deferring network upgrades, embodied carbon can be amortized over longer periods of time, resulting in more environmentally-sustainable networks.

One common reason for equipment upgrades concerns greater energy efficiency of newer generations of hardware. In many cases, such upgrades can indeed lead to a positive impact on energy usage and resulting energy bills during network operations. However, they neglect the contribution of embodied carbon into overall sustainability. To have an effect, newer,

more energy-efficient equipment needs to be in service for a certain amount of time for the operational carbon savings to offset the carbon that had been embodied in the decommissioned equipment and hence lead to a net-positive effective. This is illustrated in Figure 3. Holistic approaches to greening the network therefore need to take the entire lifecycle of network equipment into account, not only the operational phase during which pieces of network equipment are actually in use.



■ **Figure 3** Carbon emissions (C) over time (t). Carbon efficiency amortizes over time, such that there is a “sweet spot” when upgrades should happen (left). Instead, the hypes can cause more frequent upgrades, such that this sweet spot is never reached (right).

In wireless networks in particular, there are many reasons for hardware upgrades besides increasing the capacity (and *capacity* upgrades are the possibly unnecessary and hence not “green” element in this discussion). These include new abilities of modern equipment such as network slicing and being able to connect a larger number of customers in a small geographical area. In wired networks, reasons for upgrades are also manifold, as the discussion in Dagstuhl has shown:

- fiber does not have cross-talk;
- fiber is less sensitive to weather; rain affects the impedance of other wires, which makes it unusable when the snow melts;
- copper replaced by fiber can be sold back for a net profit;
- new hardware can offer new network features such as IPv6 support;
- regulatory reasons may require updates.

Addressing this begins with asking the right questions. Rather than trying to lobby against hardware upgrades, we should identify **why updates are carried out**: when are they done for capacity, and when are they done for features? It would be useful to gather more data about this (e.g., through network operator or end user surveys). Research should focus on studying (measuring, monitoring) the actual traffic demand and service to identify whether the running applications would (or would not) benefit from a capacity upgrade. This requires the development of novel measurement and monitoring tools that can enable both users and service providers to make much more informed decisions than “speed test” systems currently offer. This is in line with recent and ongoing work in the IP Performance Measurement (IPPM) Working Group of the IETF [4, 5].

References

- 1 Ramon Bister, Alexander Clemm, Severin Dellsperger, and Reto Furrer. Towards sustainable networking: Unveiling energy efficiency through hop and path efficiency indicators in computer networks. In *2024 IEEE 10th International Conference on Network Softwarization (NetSoft)*, pages 118–126, 2024.
- 2 Michael Welzl, Peyman Teymouri, Safiqul Islam, David Hutchison, and Stein Gjessing. Future internet congestion control: The diminishing feedback problem. *IEEE Communications Magazine*, 60(9):87–92, 2022.
- 3 Lloyd Brown, Yash Kothari, Akshay Narayan, Arvind Krishnamurthy, Aurojit Panda, Justine Sherry, and Scott Shenker. How I Learned To Stop Worrying About CCA Contention. In *Proceedings of the Thirty-First Workshop on Hot Topics in Networks (HotNets)*, HotNets ’23, New York, NY, USA, November 2023. Association for Computing Machinery.
- 4 Christoph Paasch, Randall Meyer, Stuart Cheshire, and Will Hawkins. Responsiveness under Working Conditions. Internet-Draft draft-ietf-ippm-responsiveness-05, Internet Engineering Task Force, October 2024. Work in Progress.
- 5 Bjørn Ivar Teigen and Magnus Olden. Quality of Outcome. Internet-Draft draft-ietf-ippm-qoo-01, Internet Engineering Task Force, September 2024. Work in Progress.

6 Talks

6.1 Energy aware orchestration and lifecycle management of Telco workloads: an NFV-MANO approach

Hammad Zafar (NEC Laboratories Europe – Heidelberg, DE, hammad.zafar@neclab.eu)

License  Creative Commons BY 4.0 International license
© Hammad Zafar

This talk highlighted the recent advancements made in the NFV-MANO orchestration framework related to energy efficiency. Furthermore, it presented a vision for how the energy efficient NFV, ‘Green NFV’, can be used in the lifecycle and control loop based energy consumption optimization of NFV-based Telco deployments. Some challenges beyond NFV-MANO and potential research directions were also highlighted.

6.2 Pathways for sustainability on transport networks


Luis Miguel Contreras Murillo (Telefónica – Madrid, ES, luismiguel.contrerasmurillo@telefonica.com)

License  Creative Commons BY 4.0 International license
© Luis Miguel Contreras Murillo

This talk focused on Transport Networks, visiting a number of possible pathways for reducing energy consumption (and optimizing other Green dimensions): (1) Reduction of network elements, (2) Rationalization on the need of cloud facilities, (3) Programmable adaptation of energy consumption, (4) Enabling the design of novel optimization algorithms, and (5) Introduction of more efficient platforms.

6.3 For a joint operator-centric approach to assessing network management effort


Marc Bruyère (IIJ Research Lab & The University of Tokyo, JP, marc@bruyere.io)

License  Creative Commons BY 4.0 International license
© Marc Bruyère

Network complexity can be overwhelming, even for experts. Changing a network system requires navigating intricate configurations and monitoring parameters to ensure seamless operations. This talk explored a novel approach to simplify this process by quantifying the complexity of different architectures. By analysing publicly available data from Internet Exchange Points (IXPs) and consulting with operators and vendors, the presented work aims to provide a practical solution for characterizing the complexity of network architecture options. The goal is to empower operators to make informed decisions and facilitate easier maintenance and upgrades.

6.4 Autopower and Network Power Zoo

Romain Jacob (ETH – Zurich, CH, jacobr@ethz.ch)

License  Creative Commons BY 4.0 International license
© Romain Jacob

This presentation briefly presented two projects related to the collection of networking power data.

“Autopower” refers to a power measurement unit designed to be deployed in a server room or datacenter; it can be remotely controlled and automatically retrieves power measurements to a server, from which it can be visualized in a web interface.

The “Network Power Zoo” is a public database designed to aggregate power-related data for networking equipment. This includes datasheet information, internal or external power measurements (such as those collected by Autopower units) or power model parameters. The network power zoo is live at: <https://networkpowerzoo.ethz.ch/>

7 Seminar outcomes

The seminar concluded with a discussion of a consolidated list of ideas that each breakout group has developed. For each idea, we assigned the potential impact, time frame and influence on grid dynamics in a plenary discussion. The outcome is shown in Tables 1, 2 and 3, which present the ideas for the breakout categories “Application, Systems, Stakeholders”, “Network Technologies” and “Lifecycle and Control Loops”, respectively.

References

- 1 Mindula Illeperuma, Rafael Pina, Varuna De Silva, and Xiaolan Liu. Novel directions for neuromorphic machine intelligence guided by functional connectivity: A review. *Machines*, 12(8), 2024.
- 2 Xiaolan Liu, Yansha Deng, and Toktam Mahmoodi. Energy efficient user scheduling for hybrid split and federated learning in wireless uav networks. In *ICC 2022 - IEEE International Conference on Communications*, pages 1–6, 2022.

- 3 Xiaolan Liu, Jiadong Yu, Yuanwei Liu, Yue Gao, Toktam Mahmoodi, Sangarapillai Lambotharan, and Danny Hin-Kwok Tsang. Distributed intelligence in wireless networks. *IEEE Open Journal of the Communications Society*, 4:1001–1039, 2023.
- 4 David Ros and Michael Welzl. Less-than-best-effort service: A survey of end-to-end approaches. *IEEE Communications Surveys & Tutorials*, 15(2):898–908, 2013.
- 5 Stanislav Shalunov, Greg Hazel, Jana Iyengar, and Mirja Kühlewind. Low Extra Delay Background Transport (LEDBAT). RFC 6817, December 2012.
- 6 Bruce Spang, Shravya Kunamalla, Renata Teixeira, Te-Yuan Huang, Grenville Armitage, Ramesh Johari, and Nick McKeown. Sammy: smoothing video traffic to be a friendly internet neighbor. In *Proceedings of the ACM SIGCOMM 2023 Conference*, ACM SIGCOMM '23, page 754–768, New York, NY, USA, 2023. Association for Computing Machinery.
- 7 Michael Welzl. Not a trade-off: On the wi-fi energy efficiency of effective internet congestion control. In *2022 17th Wireless On-Demand Network Systems and Services Conference (WONS)*, pages 1–4, 2022.

■ **Table 1** Ideas in the category “Application, Systems, Stakeholders” along with estimations of their impact, time frame of execution, and potential impact on grid dynamics.

| Idea | Impact | Time Frame | Grid dynamics |
|---|--------|------------|-------------------------|
| Applications, Systems, Stakeholders | | | |
| Trust models - increasing transparency by enabling trustworthiness: of veracity/integrity of generated/exported/aggregated data; of confidentiality of data use (i.e., for intended purpose only); What support is needed/can be provided by federation and regulation. | High | Medium | N/A |
| AI | | | |
| Carbon monitoring: Observability (measuring Joules consumptions on hardware); Tools for measuring energy / carbon emissions from training: https://codecarbon.io/ https://github.com/powerapi-ng/pyJoules | High | Now | Medium |
| More efficient ML distributed computing and transport protocols (collective communication, in-network aggregation, multi-destination delivery); ML training closer to the edge; Flexible edge intelligence (server-side training as energy is available; at the user side, energy constrained) [3, 2]. | Medium | Short | Low |
| Training: Workload adaptivity (pruning, sparsification, compression, knowledge distillation) depending on observed utilization/cost (also: mitigating traffic spikes); Focus on aggregation using historical data from previous rounds – ensure that resource limited devices can use a trained model; Can save computation energy and communication overhead. | High | Now | Low |
| Deterministic networking for training; decision making tools (e.g., use multi-agent reinforcement learning) to enable users to better select what and when to train, what inference model to use (energy impact) | Medium | Short | Medium |
| Preload models at the edge for inference: consolidate requests for AI models based on predicted / forecasted demand to a few edge nodes; offload inference to network devices: e.g. mitigating DDoS in the network can be done easily with limited cost in resources and power. | Medium | Short | Medium |
| Model: Use Neuromorphic computing, an innovative and energy-saving machine learning algorithm for the future [1]. | Medium | Long | Low |
| Power: Co-locate DC & workloads to renewable energy sources (windmills) | High | Medium | High |
| Streaming | | | |
| Reduce idle power consumption in caches (e.g. try to dynamically disable some DRAM). About Grid dynamics: idle periods of content servers typically align with the low-power consumption period. It helps to shift the overall consumption, but not shift the demand. | High | 2-10 years | See “idea” |
| Find the right balance between various network constraints and the number of cache servers needed. We need a framework to understand the optimal design for a given network topology and traffic pattern. For example, reduce cache hit rate target; right now it’s about 95% for Netflix. Maybe that’s too much? With smaller caches, we might have a net benefit in transmitting more over the network to reduce the HW costs of the caches. The trade-off to be looked at is the energy of adding more caches vs the amount of traffic that has to be carried through the network. | High | 0-1 year | N/A |
| Measure: how big is the potential energy gain from allowing devices to sleep (run fast to completion), as opposed to a “SAMMY” [6] style of pacing traffic (= more or less continuous low-rate sending)? | Medium | 0-2 years | Unclear |
| Optimize encoding parameters (e.g., segment length, luminance,...) and design energy-efficient ABR algorithms. | High | 0-2 years | Unclear |
| Find an optimal QoE / energy consumption balance at the client side. | High | 0-2 years | Unclear |
| Mobile Networks | | | |
| Continue to improve power proportionality of RAN (3GPP standards, implementations, ...) | High | 0-10 years | Applicable |
| Continue to enable shutting down entire nodes/base stations | High | 0-5 years | Applicable |
| Make virtualization in vRAN and Core more energy efficient. Identify the power consumption benefit and value for virtualizing hardware equipment in the RAN. Power proportionality should be investigated for both dedicated hardware and their virtualized counterparts. | Medium | 0-5 years | Applicable |
| Build a system view of the whole mobile network in terms of energy efficiency and carbon footprint (of sites, DCs, nodes in RAN, Edge, Core etc.) which can be useful in the overall deployment strategy of mobile network operators across the network continuum. An optimal trade-off between power consumption and other factors like latency/performance can be made by the operators based on the power/CO2 profiles of different sites and segments of the network to host different workloads like CDNs or AI/ML training/inference workloads. For example, if latency is not impacted a lot by putting a CDN a bit further from the Edge where the sites are “greener”, then it may be better to do so for overall CO2 emissions of a mobile network. | High | 0-5 years | Not applicable |
| Sharing central backhaul (but also RAN sites and RAN clouds/DCs) across multiple mobile operators could present a promising opportunity for improved efficiency, reduced costs, and enhanced sustainability. Challenges related to such multi-tenancy will need to be addressed on various layers including the virtualization/cloudification layer. Power-saving optimizations might make things more difficult as turning off the RAN/Edge sites/nodes may not be straightforward if several different tenants share the same site/nodes in the mobile network (any segment, whether RAN, Edge or Core). | High | 0-10 years | Applicable (Low-Medium) |

■ **Table 2** Ideas in the category “Network Technologies” along with estimations of their impact, time frame of execution, and potential impact on grid dynamics.

| Idea | Impact | Time Frame | Grid dynamics |
|--|--------|------------|---------------|
| Network Technology | | | |
| Routing | | | |
| Carbon transparency between ISPs, certify ISP-announced information | High | Short | N/A |
| Integrate of current protocols with carbon-aware systems such as SCION | High | Medium | Applicable |
| Agree on and standardize a set of energy metrics for network equipment | High | Short | N/A |
| Identify a mechanism to acquire trusted carbon data from energy grids | Low | Short | N/A |
| Traffic Engineering and Congestion Control | | | |
| In streaming, Less-than-best-effort (LBE) traffic [4, 5] could be used to pre-fill the client buffer more than today, switching into LBE mode. LBE traffic is more flexible, and it could be assigned to “greener” paths with worse performance if the client allows it. | Medium | Medium | Applicable |
| Quantify the reductions in QoE. Consider different traffic types and their wide area latency requirements. Game downloads or backups could tolerate 1 second or more additional delay. For Netflix, VoD doesn’t have very strict latency (in the order of seconds). Games are very different. | Medium | Medium | Applicable |
| Consumer side: energy labels could help to motivate the 20% of the conscious consumers. | High | Short | Applicable |
| Investigate how to quantize carbon information to use it as a Traffic Engineering weight | Low | Short | Applicable |
| Power Saving | | | |
| Visibility of grid-side state/information about real-time power consumption information should be used to make ICT-side optimizations. Factor in grid-side demand cycles, generated power in the system for solutions in the DC/network side. | High | Medium | Applicable |
| Matching between power side demand and ICT side workload/traffic forecasts could optimize workload scheduling and consolidation. | High | Medium | Applicable |
| Scavenging workloads can be scheduled in off-peak times at the ICT side to still keep the power system side balanced. | Medium | Medium | Applicable |
| There is a CPU-offloading trend (to smart NICs and GPUs) to meet performance requirements, e.g. of Telco workloads. These hardware accelerators are very power hungry. Can software alternatives meet the performance without being as power-intensive as hardware solutions? | Medium | Long | Low |
| Add power knobs in the backbone to disable links or do rate adaptation. | Medium | Medium | High |
| Wireless | | | |
| Continue to enable shutting down entire nodes/base stations (during nights etc.) | High | 0-10 years | High |
| Continue to improve the power proportionality of RAN (component shutdown, designing periodic transmissions better, etc.) | High | 0-5 years | High |
| Radio parameters: switch spectrum based on application (unclear if beneficial at all, but worth investigating?) | Medium | 5-10 years | Medium |
| Accommodating both quick operation to sleep more and reducing impact on other users: this paper [7] finished transfers quickly and then slept for WiFi and the “SAMMY” paper [6] paced traffic to avoid impacts to others. Can we find a common ground to both save energy and not have an impact? | Medium | 0-5 years | Low |
| Demand shifting: time dimension (cache overnight), WiFi/Wireless dimension (cache on WiFi) | Medium | 0-5 years | Medium |
| Better information sharing between applications and networks, e.g., know what bandwidth is available when, if sleeping at specific times would be helpful, etc. | Medium | 0-10 years | N/A |

■ **Table 3** Ideas in the category “Lifecycle and Control Loops” along with estimations of their impact, time frame of execution, and potential impact on grid dynamics.

| Idea | Impact | Time Frame | Grid dynamics |
|--|--------|-----------------|------------------|
| Lifecycle and Control Loops | | | |
| HW Lifecycle | | | |
| Give users and ISPs realistic measurement / monitoring tools to be able to judge: is an upgrade worthwhile? | Medium | 5 years | N/A |
| Survey operators to understand why and when they upgrade their network. | Medium | 2 years | N/A |
| “Energy Star” for networks. This is common for consumer devices and would have been subject to telco regulation in the past. There is nothing comparable for networks today, yet it would potentially be useful to make carbon-informed network upgrade decisions. | Medium | Medium | N/A |
| Observability | | | |
| Definition of metrics, balancing bottom up perspective and top-down perspective (what metrics do solutions require, what metrics do clients ask for – example, what metrics does an application provider require from an ISP to make more informed decisions). As a subaspect: Attributional vs consequential metrics – example, metrics as a share of power consumption may not be “fair”; can they help make informed decisions? | Medium | Short | N/A |
| Metrics in virtualized environments – do we need to attribute (and how) carbon metrics to virtualized instances? Associate real-time power measurements from the HW layer with VNF components (VMs, Pods) for multiple virtualized workloads sharing the same hardware | Low | Short to Medium | N/A |
| Define additional parameters to indicate attribution gaps – factors outside of the control of the individual virtual entity, such as whether or not co-hosted with other VEs. | Low | Short | N/A |
| Frequency of measurements and telemetry data – which frequency is needed to make informed decisions, synchronizing between different stakeholders (e.g. suppliers vs consumers of power) | Medium | Short | Relevant |
| Trust schemes to exchange data across vertical boundaries and horizontal boundaries that ensure privacy, integrity of aggregated data, interdomain telemetry | Medium | Medium to Long | N/A |
| Cross-organizational collaboration to optimize carbon outcomes – includes exchange of data across boundaries, support for global optimization schemes versus local schemes & assessment of local vs global optimization potential | Medium | Medium to Long | N/A |
| Address security vulnerabilities associated with observation of carbon data (e.g. false advertising to overload certain paths, etc) | Low | Medium | Applicable |
| Integration of path metrics and systems that utilize them (e.g. Carbon Aware Networking, SCION) | Medium | Long | Potentially High |
| Compliance: infrastructure/facilities for objective assessment and auditing of actual metrics | Medium | Medium | N/A |
| May need separate power monitoring infrastructure to improve accuracy, auditability, calibration – but consider its impact on the carbon equation. Inaccuracies of 15% are not uncommon today. | Low | Medium | N/A |
| Provide a library of “power profiles” | Medium | Low | N/A |
| Algorithm & Forecasting | | | |
| Tools that help users assess carbon impact of their networks | High | Medium | Relevant |
| Assess how to design prices as a way to direct towards behavior that is carbon efficient | High | Medium | Relevant |
| AI techniques to forecast traffic demand, energy consumption, etc in order to take better decisions (e.g. when to make spare capacity available, how to reconfigure routes). Assess impact of potential improvement versus increased carbon cost involved in making the forecast itself. | High | Medium | Relevant |

Participants

- Jari Arkko
Ericsson – Jorvas, FI
- Anaïs Berkes
University of Cambridge, GB
- Carsten Bormann
Universität Bremen, DE
- Marc Bruyère
Internet Initiative Japan Inc. – Tokyo, JP
- Wen Cai
University of Oslo, NO
- Alexander Clemm
Los Gatos, US
- Luis Miguel Contreras Murillo
Telefónica – Madrid, ES
- Hermann de Meer
Universität Passau, DE
- Sawsan El Zahr
University of Oxford, GB
- Simone Ferlin-Reiter
Red Hat, Stockholm, SE & Karlstad University, SE
- Te-Yuan Huang
Netflix – Los Gatos, US
- Romain Jacob
ETH Zürich, CH
- Tony John
Otto-von-Guericke-Universität Magdeburg, DE
- Dirk Kutscher
HKUST – Guangzhou, CN
- Jong Hoon Kwon
ETH Zürich, CH
- Xiaolan Liu
Loughborough University, GB
- David Oran
MIT – Cambridge, US
- Gopika Premsankar
Aalto University, FI
- Daniel Schien
University of Bristol, GB
- Magnus Skjegstad
University of Oslo, NO
- Farzad Tashtarian
Alpen-Adria-Universität Klagenfurt, AT
- Michael Welzl
University of Oslo, NO
- Cedric Westphal
University of California, Santa Cruz, US
- Jiadong Yu
HKUST – Guangzhou, CN
- Hammad Zafar
NEC Laboratories Europe – Heidelberg, DE
- Noa Zilberman
University of Oxford, GB

