Towards a power-proportional Internet

Romain Jacob
ETH Zürich

Abstract

Today, the ICT industry has a massive carbon footprint (a few percent of worldwide emissions) and one of the fastest growth rates. The Internet accounts for a large part of that footprint while being also energy inefficient; *i.e.*, the total energy cost per byte transmitted is very high. We argue that this inefficiency stems from the lack of power proportionality in computer networks and discuss some first ideas and research directions to address this issue.

1 Introduction

The Internet has become a vital component of our societies: ubiquitous, reliable, but energy inefficient [11]. Making the Internet more sustainable must be a prime goal for the coming decade. While history shows that the Internet does not change as long as "it works" [12], we argue that continuing to waste so much energy (and even more, given the expected growth) is unacceptable; **the Internet has to change**. Not pushing for research in that direction feels like a criminal omission. One can explore several directions to reduce the Internet's carbon footprint.

1. Use more sustainable energy to power the Internet;
2. Use time- and location-shifting to consume energy where and when it is the greenest;
3. Reduce the Internet energy consumption for
    - running the network, *i.e.*, communicate; that is referred to as the operational cost;
    - producing the network, *i.e.*, build and renew hardware infrastructure; that is the embodied cost.

These objectives are generally independent: harvesting and using sustainable energy is an overarching challenge; all fields and disciplines must improve in that dimension. However, **the best energy is the one we do not consume**. We must learn to use the energy that the Internet consumes better.

It is well-understood that energy efficiency results from *power proportionality*; that is, a linear scaling between the power drawn and the volume of (useful) work done. Today, the Internet is far from power proportional, but we believe there is room for improvement. In particular, we argue that we must rethink the co-design of routing protocols, networking hardware, and networking software to improve the Internet's energy efficiency without compromising performance.

2 Power-proportional Internet

Almost two decades ago, Gupta and Singh [11] presented a simple calculation showing that transiting one byte over the Internet core consumes between 2× and 24× more energy than if using wireless LAN technology. While this was a rough estimation, and some numbers are probably outdated, this ballpark comparison most likely still holds; it might even be more favorable to wireless nowadays, as energy consumption generally increases with the Ethernet bandwidth. How can it be? What makes the Internet so power-hungry?

2.1 Energy-inefficiencies of the Internet

We argue that the energy inefficiency of the Internet mainly stems from three facts.

Fact 1 Network devices are "always-on."

Fact 2 Today's networking devices are far from power proportional. Their energy usage is essentially independent of their load (Figure 1), *i.e.*, most of the cost comes from powering the device (P_0), regardless of its utilization.

Fact 3 ISP networks are largely under-utilized because (i) operators design their network for peak traffic and (ii) aggressively over-provision to avoid congestion. Operators have reported upgrading links when utilization reaches as little as 50% [5, 18].

In summary, we operate network devices most of the time close to their least energy-efficient operating point!

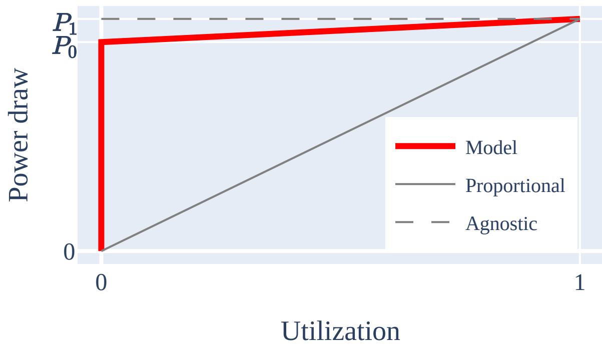

Figure 1: The typical power profile of network hardware. The largest portion of power consumption comes from powering on the hardware (P_0), which then increases only slightly with the utilization. For energy efficiency, the "ideal" power profile is proportional to the utilization. *Redrawn from [2].*

The most straightforward concept to improve power proportionality is known as *duty-cycling*; that is, keep devices off as much as possible and power them on dynamically to adjust to the traffic demand. Some research in that direction, *e.g.*, [3, 4, 13, 17], confirms there is potential to save up to 50% of energy (see § 2.3), but also highlight three practical problems when turning devices on and off often.

Convergence Turning devices on and off creates instability and convergence issues in today's routing protocols and network systems (*e.g.*, optical signal amplifiers [20]).

Management Deciding which device to power off and when is a complex optimization problem which is hard to scale to run in real-time for large networks. Moreover, it requires accurate monitoring—or predictions—of the network traffic matrix, which is challenging in itself.

Start-up time Today's routers take minutes to boot [6, 13] which makes it unfeasible to power them on quickly to adjust to the traffic load. These long boot times are explained partly by the extensive memory testing commonly performed by networking devices [8].

These are not significant problems as long as the "always-on" hypothesis holds. The Internet was designed under that assumption; energy efficiency was only a secondary objective, if at all. Some other networks were designed differently, though: *e.g.*, networked embedded systems' main requirement is energy efficiency, such as to provide long-term operations with only small batteries—or even without batteries [10].

2.2 Embedded-systems-inspired redesign

What if we redesigned Internet networks by taking inspiration from embedded systems?

Embedded systems is a field borne and grown with the mindset of managing energy scarcity. In 1999, the field was ushered with the vision of the "Smart Dust" [15], a large network of tiny devices embedding a small battery, some sensors, a solar panel to harvest energy, and a wireless transmitter and receiver; all that being highly energy-efficient to sustain autonomous operation. Two decades later, technology is getting ever closer to fulfilling that vision, thanks to:

Progress in hardware *e.g.*, non-volatile memory allows devices to maintain state over powered-off periods.

Progress in software *e.g.*, custom operating systems using less state, booting efficiently, and featuring multiple modes of operation to save power.

Progress in protocol design *e.g.*, reliable end-to-end communication over unreliable wireless links.

For wireless embedded systems, the radio is the largest energy consumer.[1] Thus, energy efficiency dictates turning the radio off for as much as possible—which is somewhat comparable to turning networking devices off (§ 2.1). Low-power wireless networking became a research field aiming to communicate efficiently between devices that are often "off" from the network's perspective. Many protocols have been designed, from fully centralized to distributed ones, with various cost-benefits trade-offs.

The protocol is (or should be) the starting point for an efficient system design, as it defines the resource requirements (*e.g.*, memory, compute) and the software abstractions needed to implement it. As system researchers, we are often constrained by the existing hardware and software stacks; we adapt our protocols or algorithms to be implementable on today's devices. These constraints fundamentally limit the achievable performance compared to a clean-slate redesign of protocol, hardware, and software.

To address the energy challenge in Internet networks, we argue that we must set these constraints aside for a moment. Instead of tweaking traditional routing protocols, as attempted in previous works [3,11], we argue that improving the Internet energy efficiency is achievable with new protocols—designed under the assumption that devices will frequently turn themselves off—by taking inspiration from the vast literature on networked embedded systems. Naturally, it presents different design challenges and opportunities since the wireless and wireline physical layers are different. These new protocols will then lay out the practical challenges we must address on the hardware and software sides (see § 2.4).

[1] It was historically the case. Things are starting to change with transceivers getting more and more low-power and the push for more computing at the edge. But the point remains.

2.3 How much is there to gain?

In a perfect world, network devices could be powered on/off instantaneously, and running the network at 100% utilization does not result in congestion or packet losses. Under these assumptions and considering the power model in Figure 1, we derive and compare the energy consumed by

- running a network at 100% utilization for short periods and turning all devices off the rest of the time;
- running the network at a baseline utilization U, always.

The potential savings depend on the baseline utilization U, as well as the P_0/P_1 ratio in the power model (see Figure 1). More precisely, the power draw $P(U)$ is given by

$$P(U) = P_0 + (P_1 - P_0) \cdot U \tag{1}$$

and the time taken to transmit B bytes on a link of capacity C with utilization U is

$$t(B,C,U) = B/(C \cdot U) \tag{2}$$

We can then compute the energy E_U and $E_{U_{max}}$ consumed to transmit at utilization U and $U_{max} = 1$, respectively, and the resulting energy savings $S = (E_U - E_{U_{max}})/E_U$.

$$\begin{aligned} E_U &= P(U) \cdot t(B,C,U) \\ &= B/C \cdot (P_0/U + P_1 - P_0) \end{aligned} \tag{3}$$

$$\begin{aligned} E_{U_{max}} &= P_1 \cdot t(B,C,1) \\ &= P_1 \cdot B/C \end{aligned} \tag{4}$$

$$\Rightarrow \qquad S = \frac{P_0 \cdot (1-U)}{P_0 + (P_1 - P_0) \cdot U} \tag{5}$$

Figure 2 shows the savings from Eq. (5) with $P_1 = 1$ and different values for P_0. As expected, the smaller the baseline utilization, the more potential savings; if traffic demands an average utilization of 99%, there is little margin to turn devices off. However, given the state of practice in network capacity planning [5, 18], we rather expect the average utilization in ISP networks today to be in the low tens of percent. Moreover, the larger P_0/P_1, the more potential savings; according to the literature, we expect this ratio to be at most 0.5 for standard devices [2, 4, 13]. Assuming a baseline utilization below 30%, this yields $\geq 50\%$ energy savings!

Naturally, this is a rough approximation of the potential savings, but it provides a useful upper bound of what one can hope to achieve. It also highlights some practical research challenges to realize such savings, *e.g.*, fast powering-on/off times and running the network reliably at high utilization.

2.4 Practical challenges

Given the power profile of today's network devices (Figure 1), it is clear that turning them off whenever possible[2] has some

[2]Or turning off only certain components [4]

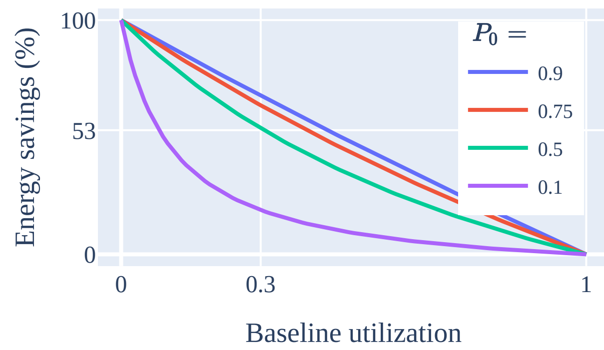

Figure 2: Even assuming that powering up a device accounts for only 50% of its maximum power draw ($P_0 = 0.5$), the potential energy savings are larger than 50% for baseline utilization up to 30%. *Eq. (5) with* $P_1 = 1$.

potential to reduce the Internet operational costs significantly (§ 2.3). However, achieving this potential requires reducing the start-up time of networking devices.

More precisely, to harness the energy benefits without dramatically increasing delay, we must bring the time for a device to go from a low-power mode to "ready to forward traffic" down from the *minute* scale to the *millisecond* scale. We believe this is possible, but it will demand redesigning the networking hardware, software, and protocols. Here are some of the directions we are exploring:

Redesign routing protocols

- to be more distributed and cope better with network nodes switching on/off often;
- to require less state in network devices, hence requiring less memory;[3]
- to tolerate transient faults resulting, *e.g.*, of using an approximate forwarding state.

Redesign networking hardware

- to include non-volatile memory and speed up the state reconstruction after powering on;
- to improve power proportionality by power-gating peripherals, using DVFS[4], or running RAPL.[5] [6]

Redesign networking software

- to extend wake-on-LAN [1] to networking devices;
- to optimize the boot time of network devices' OS.

It is an ambitious plan, but the experience of embedded systems has shown that co-designing hardware, software, and protocols can push the energy efficiency of the overall system design very far. We strongly believe that the wireline networking community must follow that example and redesign the Internet to make it more energy-efficient.

[3]*E.g.*, using source-based routing mechanisms as explored in [14]

[4]Dynamic Voltage and Frequency Scaling

[5]Running Average Power Limit

[6]These technologies are now common on servers. Could they be transferred to networking devices, and how much is there to gain with those?

3 Research directions

We foresee several research directions toward a more power-proportional Internet, some of which we have already started exploring. The IAB workshop would be an excellent opportunity to discuss those ideas further.

Several of these ideas stemmed from discussions during the Dagstuhl seminar on Power and Energy-aware Computing on Heterogeneous Systems (PEACHES – August 2022) [19].

The potential for sleeping in networks ISP networks often observe strong daily patterns in their traffic. Even with today's long boot times and ill-suited protocols, turning off some routers or switches on a daily timescale would be feasible. But how much would there be to gain?
➜ We are analyzing existing network topologies to assess how many links and nodes one can turn off while maintaining connectivity.

Where does power go? It is unclear what the main power consumer elements in today's networking devices are. Existing studies are dated and imprecise. With a better understanding of which hardware elements draws the most power, can we redesign network devices for better power proportionality?
➜ We started a power-profiling study of networking devices to understand how device utilization affects energy consumption. Essentially, we want to instantiate the model shown in Figure 1. We are starting with programmable switches looking at the impact of the programmable switch logic, the number of pipeline stages used, the memory usage, port utilization, etc.

Power proportionality at the physical layer Optical links make it technically feasible to turn on/off some wavelengths to consolidate traffic on fewer ones. What would be the consequences on traffic if we were doing this (very) often? Would any adaption of the routing protocol be required? Most importantly, could this yield significant energy savings?
➜ Some recent related works on optical wavelength management [16, 20] should be explored.

Graceful degradation of service Every network has to turn off or reboot some devices from time to time for various maintenance tasks. There is extensive literature on "graceful maintenance" (*e.g.*, [7, 9]), but what is the state of practice today? How do operators "prepare" their network for such planned interruptions of service? How frequently are those happening? How long does it take in practice to turn things on and off today?
➜ We know that shutting devices off (when there are not strictly needed) is an efficient energy-saving vector. Hence, if one does it routinely for network updates, could we do it more frequently for energy-saving purposes?

References

[1] AMD. Magic Packet Technology, November 2015. URL: https://www.amd.com/system/files/TechDocs/20213.pdf.

[2] Aruna Prem Bianzino, Claude Chaudet, Federico Larroca, Dario Rossi, and Jean-Louis Rougier. Energy-aware routing: A reality check. In *2010 IEEE Globecom Workshops*, pages 1422–1427, December 2010. doi:10.1109/GLOCOMW.2010.5700172.

[3] Aruna Prem Bianzino, Luca Chiaraviglio, Marco Mellia, and Jean-Louis Rougier. GRiDA: GReen Distributed Algorithm for energy-efficient IP backbone networks. *Computer Networks*, 56(14):3219–3232, September 2012. URL: https://www.sciencedirect.com/science/article/pii/S1389128612002344, doi:10.1016/j.comnet.2012.06.011.

[4] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, and S. Wright. Power Awareness in Network Design and Routing. In *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, pages 457–465, April 2008. doi:10.1109/INFOCOM.2008.93.

[5] Cisco. Cisco WAN Automation Engine (WAE) - Best Practices in Core Network Capacity Planning White Paper. URL: https://www.cisco.com/c/en/us/products/collateral/routers/wan-automation-engine/white_paper_c11-728551.html.

[6] Cisco. Cisco 3750, 2900, 3900 and 4400 Boot Time, November 2018. URL: https://community.cisco.com/t5/switching/cisco-3750-2900-3900-and-4400-boot-time/td-p/3739176.

[7] Francois Clad, Stefano Vissicchio, Pascal Merindol, Pierre Francois, and Jean-Jacques Pansiot. Computing Minimal Update Sequences for Graceful Router-Wide Reconfigurations. *IEEE/ACM Transactions on Networking*, 23(5):1373–1386, October 2015. URL: http://ieeexplore.ieee.org/document/6850090/, doi:10.1109/TNET.2014.2332101.

[8] Dell. POST and Boot Processes. URL: https://www.dell.com/support/kbdoc/en-uk/000128270/post-and-boot-procedures?lwp=rt.

[9] Pierre Francois. *Improving the Convergence of IP Routing Protocols*. Doctoral Thesis, 2007.

[10] Kai Geissdoerfer and Marco Zimmerling. Learning to Communicate Effectively Between Battery-free Devices. In *19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*, pages 419–435,

2022. URL: https://www.usenix.org/conference/nsdi22/presentation/geissdoerfer.

[11] Maruti Gupta and Suresh Singh. Greening of the internet. In *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '03, pages 19–26, New York, NY, USA, August 2003. Association for Computing Machinery. doi:10.1145/863955.863959.

[12] M. Handley. Why the Internet only just works. *BT Technology Journal*, 24(3):119–129, July 2006. doi:10.1007/s10550-006-0084-z.

[13] Brandon Heller, Srini Seetharaman, Priya Mahadevan, Yiannis Yiakoumis, Puneet Sharma, Sujata Banerjee, and Nick McKeown. ElasticTree: Saving energy in data center networks. In *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, NSDI'10, page 17, USA, April 2010. USENIX Association. URL: https://www.usenix.org/legacy/event/nsdi10/tech/full_papers/heller.pdf.

[14] Seng-Kyoun Jo, Lin Wang, Jussi Kangasharju, and Max Mülhäuser. Eco-friendly Caching and Forwarding in Named Data Networking. In *2020 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN*, pages 1–6, July 2020. doi:10.1109/LANMAN49260.2020.9153230.

[15] J. M. Kahn, R. H. Katz, and K. S. J. Pister. Next century challenges: Mobile networking for"Smart Dust". In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, MobiCom '99, pages 271–278, New York, NY, USA, August 1999. Association for Computing Machinery. doi:10.1145/313451.313558.

[16] Mehrdad Khani, Manya Ghobadi, Mohammad Alizadeh, Ziyi Zhu, Madeleine Glick, Keren Bergman, Amin Vahdat, Benjamin Klenk, and Eiman Ebrahimi. SiP-ML: High-bandwidth optical network interconnects for machine learning training. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*, SIGCOMM '21, pages 657–675, New York, NY, USA, August 2021. Association for Computing Machinery. doi:10.1145/3452296.3472900.

[17] Sergiu Nedevschi, Lucian Popa, Gianluca Iannaccone, Sylvia Ratnasamy, and David Wetherall. Reducing Network Energy Consumption via Sleeping and Rate-Adaptation. In *5th USENIX Symposium on Networked Systems Design and Implementation (NSDI 08)*, 2008. URL: https://www.usenix.org/conference/nsdi-08/reducing-network-energy-consumption-sleeping-and-rate-adaptation.

[18] Hank Nussbacher. Nanog: Bottlenecks and link upgrades. URL: https://seclists.org/nanog/2020/Aug/193.

[19] Schloss Dagstuhl. Dagstuhl Seminar on Power and Energy-aware Computing on Heterogeneous Systems (PEACHES). URL: https://www.dagstuhl.de/en/program/calendar/semhp/?semnr=22341.

[20] Zhizhen Zhong, Manya Ghobadi, Maximilian Balandat, Sanjeevkumar Katti, Abbas Kazerouni, Jonathan Leach, Mark McKillop, and Ying Zhang. BOW: First Real-World Demonstration of a Bayesian Optimization System for Wavelength Reconfiguration. In *Optical Fiber Communication Conference (OFC) 2021*, page F3B.1, Washington, DC, 2021. OSA. URL: https://opg.optica.org/abstract.cfm?URI=OFC-2021-F3B.1, doi:10.1364/OFC.2021.F3B.1.