Research Avenues Towards Net-Zero Cloud Platforms

Daniel S. Berger, Fiodar Kazhamiaka, Esha Choukse, Íñigo Goiri, Celine Irvene, Pulkit A. Misra,
Alok Kumbhare, Rodrigo Fonseca, Ricardo Bianchini   (Microsoft Azure)

Introduction The carbon impact of cloud providers spans on-site generators and staff emissions (scope 1), purchased energy (scope 2), and carbon embedded in chips, PCBs, racks/server enclosures, and buildings (scope 3). Scope 1 contributes negligible amounts. Before shifting to renewable energy (avenue 2 below), scope 2 made up the majority of carbon emissions. With renewable energy, scope 3 dominates carbon emissions and cloud providers pursue reducing and then offsetting the remaining emissions [2, 5, 13]. With the high cost of carbon offsets, reducing carbon emissions will become a long-standing priority.

This abstract describes six high-impact reseach avenues that can reduce the carbon emissions of cloud platforms by an order of magnitude. While based on internal life cycle assessments, they describe only a subset of active initiatives at Microsoft. Our carbon impact estimates are subject to confidentiality agreements with third parties to evaluate scope 3 impacts and they may change as life cycle assessments become increasingly detailed. Consequently, we focus on a qualitative description of sustainability avenues.

1) Resource Efficiency Improving resource utilization uniformly improves efficiency across all scopes. In the last decade, cloud providers have made significant progress in their efficiency efforts. At Azure, more than 70% of CPU cores have virtual machines (VMs) scheduled on them [6]. Key enablers are reducing server pools reserved only for some engineering teams (resource silos), scheduling improvements, and harvesting unallocated resources [1,6]. However, scheduling a VM on a core does not mean the core is actually used. Roughly $\frac{3}{4}$ of VMs have less than 25% CPU utilization [3] as cloud customers overprovision VMs and convincing customers to rightsize VMs is not always successful. We need systems that oversubscribe allocated-but-unused resources to close the utilization gap. A key research challenge remains VM opaqueness, which requires fusing cloud-scale and node-level telemetry [15].

2) Energy Efficiency Cloud servers and accelerators allow trading-off performance and energy efficiency using multiple modes of operation. Software can further optimize for performance per watt using fine-grained knobs for transitioning between the processor P and C states. While deeper C-states in processors allow better power savings (up to 50% lower idle power in some cases), it takes longer to transition out of them into active mode [16]. A key challenge is that cloud datacenter always must be ready to respond to sudden traffic with minimal latency, making the trade off between energy and performance difficult. Offering multiple tiers of performance can enable cloud providers to use these features better.

3) Renewable Energy Running cloud datacenters on renewable energy will eliminate almost all scope 2 carbon. Google has been purchasing 100% renewable energy since 2017 [5] with AWS [2] and Microsoft [13] reaching 100% in 2025. Importantly, these commitments rely on non-temporal and non-spatial power purchase agreements, which means that they are not carbon free as datacenters will continue to consume energy during hours of the day and in locations when no renewable energy is actually available. Google and Microsoft are pursuing temporal (hourly) matching by 2030 [9], and spatial (datacenter location) matching is also in the works. Temporal and spatial matching of power supply requires over-provisioning or storage which drive up cost asymptotically as the matched energy percentage approaches 100% [9]. This cost curve will create significant incentives for both providers and customers, e.g., to shift demand temporally and spatially by leveraging workload flexibility. Fortunately, shifting even a small fraction of workloads can be effective due to the asymptotic cost curve. However, enabling carbon-aware demand shifting at scale remains open.

4) Server and Network Lifetimes Compute servers, storage servers, and the network are the main contributors to the remaining scope 3 emissions. Buildings, power, and cooling infrastructure are smaller contributors due to longer lifetimes. Thus, cloud providers seek to extend the lifetime of servers and network components. This reduces carbon by amortizing their carbon cost over time and delaying the manufacturing of new servers. Generally, the cloud industry is already moving in this direction with Microsoft recently increasing the depreciable lifetime of servers and network equipment from four to six years [7]. Depreciable lifetime is a conservative lower bound used in financial accounting. In practice, servers are deployed far longer than their depreciable lifetime. Thus, research should focus on extending lifetime well beyond a dozen years to push the envelope compared to existing industry behavior. Extremely-long lifetimes introduce multiple challenges. First, while we see little indication of increasing failure rates after eight years, procuring parts to effectively repair servers becomes difficult. We thus need to find ways to effectively reduce repair needs. Second, workload scheduling has to account for the resulting heterogeneous performance. Interestingly, customers continue to use VM types introduced for older hardware generations. Thus, there is natural demand

for older servers. However, this also motivates the need to develop customer tools that enable cloud customers to navigate VM types. Finally, even if performance demands can be met on older servers, they may still suffer from lower cores-per-watt efficiency [14] which may make it cost-prohibitive to continue running them for a dozen years. Research on energy and carbon breakpoints for server upgrades [10] is needed.

5) Datacenter Design Novel datacenter environments better support long server lifetimes. For example, keeping temperatures stable and preventing oxidization reduced component failures rates by 88% in Project Natick [4]. At Azure, we are exploring how to achieve this environment at scale using liquid immersion cooling [12]. Besides reduced failure rates, immersion also improves power usage effectiveness and enables better utilizing deployed hardware by overclocking [8]. Power usage and hardware utilization matter particularly in the context of operating costs due to long lifetime and renewable energy. Immersion cooling also eliminates water use that is sometimes used to cool current datacenter generations.

6) Redesigning Server Infrastructure Finally, we need to focus on specific components that drive up embedded carbon. Specifically, DRAM is the dominant contributor to server carbon cost due to the high number of DRAM chips in each server (24-32 DIMMs in an Olympus-generation server). Recent studies suggest that cloud providers can move 40% of DRAM into CXL-based pools with negligible impact on performance [11]. Pooling reduces stranding and thus cloud providers need to buy less DRAM. More importantly, CXL enables cloud providers to reuse DDR4 in their DDR5 servers, and DDR5 in their DDR6 servers. Reusing DDR4 DIMMs from decommissioned servers is very effective at reducing embedded carbon costs. However, to scale up this approach, we need to find ways to increase the percentage of reused server memory, which requires research to advance performance management, e.g., by enabling dynamically moving memory between CXL and local DRAM.

Discussion These avenues bring an exciting breadth of research questions. Additionally, we hope to start a discussion on characterizing cloud carbon footprints after progressing along these avenues. Which components will dominate a cloud's carbon footprint in the future and how could we address them? How will cloud workloads and architectures shift and invalidate assumptions that are based on what we see today?

References

[1] Pradeep Ambati, Íñigo Goiri, Felipe Frujeri, Alper Gun, Ke Wang, Brian Dolan, Brian Corell, et al. Providing slos for resource-harvesting vms in cloud platforms. In *USENIX OSDI*, pages 735–751, 2020.

[2] AWS. Energy transition. https://aws.amazon.com/energy/sustainability/ accessed 1/15/2023, 2021.

[3] Eli Cortez, Anand Bonde, Alexandre Muzio, Mark Russinovich, Marcus Fontoura, and Ricardo Bianchini. Resource central: Understanding and predicting workloads for improved resource management in large cloud platforms. In *ACM SOSP*, pages 153–167, 2017.

[4] Ben Cutler, Spencer Fowers, Eric Peterson, and Mike Shepperd. Project natick. https://natick.research.microsoft.com/ accessed 1/15/2023, 2020.

[5] Google. 24/7 clean energy – data centers. https://www.google.com/about/datacenters/cleanenergy/ accessed 1/15/2023, 2021.

[6] Ori Hadary, Luke Marshall, Ishai Menache, Abhisek Pan, Esaias E Greeff, David Dion, et al. Protean:vm allocation service at scale. In *USENIX OSDI*, pages 845–861, 2020.

[7] Amy Hood. Q4 2022 earnings call, July 2022.

[8] ACM/IEEE ISCA. Cost-efficient overclocking in immersion-cooled datacenters. pages 623–636, 2021.

[9] Lucas Joppa. Made to measure: Sustainability commitment progress and updates. https://blogs.microsoft.com/blog/2021/07/14/made-to-measure-sustainability-commitment-progress-and-updat accessed 1/15/2023, 2021.

[10] Donald Kline, Nikolas Parshook, Xiaoyu Ge, Erik Brunvand, Rami Melhem, Panos K. Chrysanthis, and Alex K. Jones. Greenchip: A tool for evaluating holistic sustainability of modern computing systems. *Sustainable Computing: Informatics and Systems*, 22:322–332, 2019.

[11] Huaicheng Li, Daniel S Berger, Lisa Hsu, Daniel Ernst, Pantea Zardoshti, et al. Pond: Cxl-based memory pooling systems for cloud platforms. In *ACM ASPLOS*, 2023.

[12] John Roach. To cool datacenter servers, microsoft turns to boiling liquid. https://news.microsoft.com/source/features/innovation/datacenter-liquid-cooling/ accessed 1/15/2023, April 2021.

[13] Brad Smith. Microsoft will be carbon negative by 2030. https://blogs.microsoft.com/blog/2020/01/16/microsoft-will-be-carbon-negative-by-2030/ accessed 1/15/2023, 2020.

[14] Amanda Tomlinson and George Porter. Something old, something new: Extending the life of cpus in datacenters. In *HotCarbon*, July 2022.

[15] Yawen Wang, Daniel Crankshaw, Neeraja J Yadwadkar, Daniel Berger, Christos Kozyrakis, and Ricardo Bianchini. Sol: safe on-node learning in cloud platforms. In *ACM ASPLOS*, pages 622–634, 2022.

[16] Jawad Haj Yahya, Haris Volos, Davide B. Bartolini, Georgia Antoniou, Jeremie S. Kim, Zhe Wang, Kleovoulos Kalaitzidis, Tom Rollet, et al. Agilewatts: An energy-efficient cpu core idle-state architecture for latency-sensitive server applications. In *IEEE/ACM MICRO*, pages 835–850, 2022.