Toward Sustainable Computing with FPGAs and Chiplets

Chetan Choppali Sudarshan
Arizona State University
Tempe, USA
cchoppal@asu.edu

Aman Arora
Arizona State University
Tempe, AZ, USA
aman.kbm@asu.edu

Vidya Chhabria
Arizona State University
Tempe, AZ, USA
vachhabr@asu.edu

Abstract

The unprecedented growth of artificial intelligence (AI) and large-scale computing systems has significantly increased the environmental impact of modern computing infrastructure. Sustainable computing requires a holistic view of system design where both embodied carbon footprint (CFP), arising from design, manufacturing, packaging, and end-of-life processing, and operational CFP, arising during system deployment and use, are jointly considered. As Moore's law slows, computing systems are increasingly shifting toward heterogeneous integration (HI) architectures that combine specialized chiplets within advanced packages to improve performance, energy efficiency, scalability, and modularity. HI systems with FPGA chiplets represent an important pathway toward sustainable computing by combining modular chiplet-based integration with FPGA reconfigurability and adaptability. Such systems can potentially reduce lifecycle CFP while continuing to provide strong performance, flexibility, and support for rapidly evolving AI workloads. This paper discusses modeling and optimization approaches for HI systems with FPGA chiplets with CFP treated as a primary optimization metric alongside traditional design objectives. The paper further highlights research challenges spanning architecture, packaging, compilers, runtime systems, manufacturing, and EDA, and outlines future opportunities for sustainability-aware design.

ACM Reference Format:
Chetan Choppali Sudarshan, Aman Arora, and Vidya Chhabria. 2026. Toward Sustainable Computing with FPGAs and Chiplets. In *International Green and Sustainable Computing Conference (IGSC 2026), June 22–24, 2026, Canandaigua, NY, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3797248.3816188

1 Introduction

Modern computing, from edge devices to cloud data centers, has a substantial carbon footprint (CFP) [9, 25]. The rapid growth of artificial intelligence (AI), cloud services, and edge computing has driven an unprecedented demand for computation, making the information and communications technology (ICT) sector a meaningful contributor to global greenhouse gas emissions [7]. For decades, semiconductor innovation focused primarily on improving performance, power, and area, with sustainability largely framed through operational energy efficiency. While reducing power during use remains important, it is now clear that this view is incomplete.

The total CFP of a computing system spans its full lifecycle, including design, manufacturing, packaging, testing, deployment, operation, and end-of-life (EOL) [28, 39]. In many emerging systems, especially highly specialized or short-lived devices, these embodied emissions can rival or even exceed those incurred during operation. The relative dominance of embodied versus operational CFP remains an active and debated topic in the community, further underscoring the need for accurate modeling of both components rather than optimizing for one in isolation [1, 3]. This calls for a broader carbon-aware view of computer system design. Rather than treating sustainability as a downstream reporting exercise, carbon must be elevated to a first-order design objective alongside traditional metrics such as performance, power, area, and cost (PPAC) [6, 9, 12, 25]. Existing work has shown that lifecycle-aware modeling can expose important tradeoffs across technology choices, manufacturing decisions, and system architectures [27].

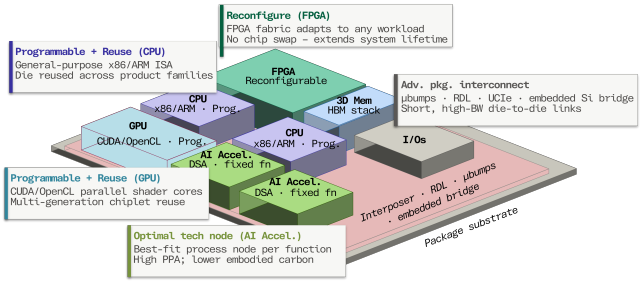


Figure 1: HI system with FPGA chiplets highlighting potential as a sustainable computing platform.

At the same time, the architectures of computing systems are changing. The slowdown of Moore's law, increasing reticle constraints, rising non-recurring engineering costs, and the growing diversity of workloads have all pushed the industry away from the assumption that one large monolithic die is always the best solution [10, 23]. Heterogeneous integration (HI) and chiplet-based design have emerged as compelling alternatives, enabling systems to be composed from smaller dies connected through advanced packaging technologies [10]. Chiplets offer several advantages that are directly relevant to sustainability: smaller dies can improve yield, reduce silicon waste, enable reuse of pre-designed components, and allow designers to mix technology nodes more effectively [25, 38].

Reconfigurable computing adds another important dimension to today's computing. Field-programmable gate arrays (FPGAs) occupy a unique position in the design space between the general-purpose processors and custom application-specific integrated circuits (ASICs). FPGAs are increasingly used as accelerators for rapidly evolving workloads such as AI/ML, networking, signal processing, and edge computing. Their programmability enables deployment-specific customization and adaptation to changing algorithms, models, and workload characteristics without requiring new hardware fabrication. This flexibility is particularly valuable for modern AI workloads, where model architectures, precision requirements, sparsity techniques, and dataflow patterns continue to evolve rapidly

This work is licensed under a Creative Commons Attribution 4.0 International License.
IGSC 2026, Canandaigua, NY, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2520-3/26/06
https://doi.org/10.1145/3797248.3816188

after system deployment. From a sustainability perspective, unlike ASICs, FPGAs can be repurposed across applications and updated after deployment, allowing embodied emissions to be amortized over multiple use cases and longer effective lifetimes [26, 28]. FPGAs raise the possibility that flexibility itself can be a carbon-saving mechanism when viewed over the complete lifecycle [26].

The combination of HI and FPGA chiplets as shown in Fig. 1 creates a particularly compelling pathway toward sustainable computing. These systems integrate reconfigurable FPGA fabrics alongside reusable chiplets such as memory, networking, connectivity, and accelerators within a common package. Such systems create new tradeoffs among efficiency, flexibility, reuse, package complexity, manufacturability, and lifecycle sustainability. On one side, chiplet-based decomposition and FPGA-based reconfigurability provide sustainability advantages through improved yield, selective technology-node assignment, reuse across deployments, and amortization of embodied CFP across multiple applications and longer system lifetimes. On the other side, advanced packaging technologies can improve bandwidth, reduce communication overheads, and enable tighter integration, while FPGA chiplets can serve as adaptable accelerators for rapidly evolving workloads, helping maintain competitive PPA characteristics despite the packaging overheads.

Realizing the sustainability potential of HI systems with FPGA chiplets requires modeling, optimization, and design methodologies that treat both embodied and operational CFP as a first-class optimization metric. Sustainability is therefore no longer only a deployment concern, but increasingly a design-space exploration and optimization problem that spans the entire computing stack. In this paper, we present a perspective on HI systems with FPGA chiplets as an important pathway toward sustainable computing. The key contributions of this paper are as follows:

- We discuss the sustainability advantages and trade-offs of HI, FPGA-based systems, and HI systems with FPGA chiplets, highlighting modularity, reuse, and reconfigurability.
- We highlight CFP modeling and optimization frameworks for HI systems and FPGA systems.
- We discuss how sustainability introduces new optimization challenges in architecture exploration, chiplet partitioning, compilation, runtime adaptation, and physical design.
- We outline research opportunities across the design stack to enable sustainable HI systems with FPGA chiplets, including future directions in broader environmental sustainability metrics, such as water footprint.

2 Sustainable Computing Platforms

2.1 HI for sustainable computing

As the semiconductor industry moves beyond large monolithic designs, HI is emerging as an important pathway for sustainable computing. In HI, systems are composed from multiple smaller chiplets integrated within a common package, enabling flexible partitioning across technology nodes, vendors, and design styles. This modularity creates opportunities to jointly optimize performance, manufacturability, reuse, and sustainability.

Yield: A primary sustainability advantage of HI is improved manufacturing yield. As shown in Fig. 2, increasing die area increases CFP while reducing yield, particularly near reticle-scale limits. By

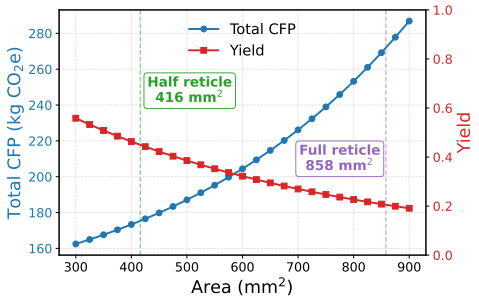


Figure 2: Variation in CFP and yield with die area for 7nm [25], with representative half-reticle and full-reticle limits marked [20].

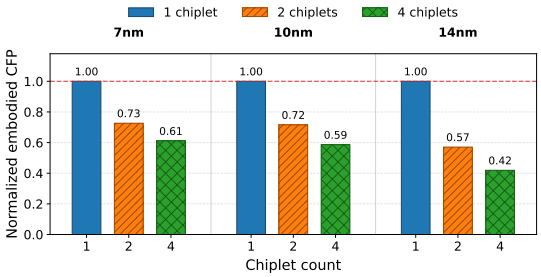


Figure 3: Normalized embodied CFP for a 550 mm^2 design using RDL-based integration across 7nm, 10nm, and 14nm nodes [25].

partitioning large monolithic designs into smaller chiplets, HI can improve die-level yield and reduce manufacturing-related embodied CFP. However, these benefits depend on package-level yield and assembly overheads.

Reuse and mix-and-match: HI also enables chiplet reuse and sustainability-aware technology-node assignment. Reusable chiplets amortize design and verification effort across multiple systems and product generations. In addition, HI enables different chiplets to be fabricated in different technology nodes depending on functionality.

Fig. 3 shows that disaggregating a 550 mm^2 design into multiple chiplets reduces normalized embodied CFP across 7nm, 10nm, and 14nm technologies, with larger reductions observed in mature nodes due to improved yield and lower process carbon intensity. HI therefore enables compute-intensive chiplets to leverage advanced nodes, while memory, I/O, analog, and connectivity chiplets can remain in mature technologies where scaling provides limited benefit. However, aggressive disaggregation also increases package-related CFP. Fig. 4 shows that progressively disaggregating a monolithic 600 mm^2 design into multiple chiplets initially reduces total embodied CFP due to yield improvements, but advanced packaging CFP becomes increasingly dominant as chiplet count grows. This demonstrates that the most sustainable HI design is not necessarily the one with the highest degree of disaggregation, but rather one that balances die-level yield benefits against package-level overheads.

2.2 FPGAs as sustainable computing solutions

Reconfiguration is an important sustainability advantage of programmable platforms because the same device can be reprogrammed across multiple applications by loading new hardware configurations, unlike an ASIC, which typically requires a separate hardware

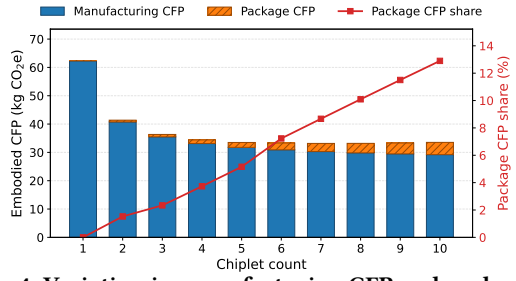


Figure 4: Variation in manufacturing CFP and package CFP with progressive disaggregation of a monolithic chip into chiplets with EMIB interconnections [25].

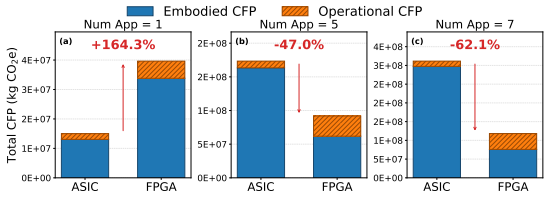


Figure 5: Total CFP comparison between ASIC and FPGA for (a) Num App = 1, (b) Num App = 5, and (c) Num App = 7, showing embodied and operational CFP components [26].

instance for each new application. This flexibility enables the embodied CFP of the FPGA to be amortized across multiple workloads and over a longer effective lifetime [4, 28]. Fig. 5 illustrates this amortization effect by comparing the total CFP of FPGA- and ASIC-based implementations as the number of applications increases. The comparison is performed at iso-performance, so the ASIC and FPGA implementations are evaluated at the same target performance while accounting for their respective power and area. [30]. For a single application, the FPGA has a higher total CFP because its larger area at iso-performance leads to higher embodied and operational contributions. However, as the number of applications grows, the FPGA can be reconfigured and reused, so its embodied CFP is incurred once and amortized across applications, whereas the ASIC typically requires a distinct hardware instance for each application. Consequently, the total CFP advantage shifts toward the FPGA, as seen in the plot for $N_{\text{app}} = 5$ and $N_{\text{app}} = 7$, highlighting reconfigurability as a key pathway to lower lifecycle CFP in multi-application settings.

2.3 HI systems with FPGA chiplets

By combining the sustainability benefits from HI and FPGAs, an HI system with FPGA chiplets packaged together with other CPU/GPU chiplets and specialized chiplets, such as networking chiplets, connectivity chiplets, high-speed transceivers, memory chiplets, custom ASIC, and other domain-specific components, can serve as a strong candidate for sustainable computing. HI with FPGA chiplets is not simply the sum of chiplet modularity and FPGA reconfigurability. It creates a new sustainability design space in which chiplet partitioning, package integration, FPGA fabric sizing, runtime mapping, and workload reuse jointly determine the lifecycle CFP. In this system, the FPGA continues to provide reconfigurability and workload adaptability, while the surrounding chiplets with advanced packaging provide specialized capabilities that improve bandwidth, connectivity, acceleration, and energy efficiency. In CPU/GPU-based HI systems, FPGA chiplets provide a programmable adaptation layer for post-deployment workload changes. By enabling post-deployment customization and adaptation, FPGA chiplets can extend the useful lifetime of the overall HI system, reducing the need for frequent hardware replacement.

This growing complexity creates a strong need for accurate modeling and optimization methodologies for HI systems with FPGA chiplets. In particular, lifecycle-aware models are needed to capture the impact of design decisions on embodied and operational CFP across manufacturing, packaging, deployment, reuse, and end-of-life stages. These models must then be integrated into optimization and design-space exploration frameworks that jointly optimize sustainability alongside traditional metrics such as PPA.

3 Carbon footprint modeling

The CFP of an HI system with FPGA chiplets is best understood over its full lifecycle, not only during its deployment phase. Lifecycle CFP is typically divided into two components: *embodied carbon*, which captures emissions incurred during design, manufacturing, testing, packaging, integration, and end-of-life processing, and *operational carbon*, which captures emissions generated during system operation. Emissions from chiplet fabrication, advanced packaging, assembly, and system integration can remain substantial and, in some cases, dominate the total CFP [9, 25–27]. The total lifecycle CFP of HI systems with FPGA chiplets serving multiple applications can therefore be expressed as the sum of embodied and deployment CFP contributions across all deployed system instances:

$$C_{\text{total}} = \sum_{k=1}^{N_{\text{sys}}} \left(C_{\text{emb},k} + \sum_{i=1}^{N_{\text{app},k}} C_{\text{deploy},k,i} \right) \tag{1}$$

where N_{sys} denotes the number of deployed HI system instances and $N_{\text{app},k}$ denotes the number of applications executed on system instance k. $C_{\text{emb},k}$ represents the embodied CFP of HI system instance k, while $C_{\text{deploy},k,i}$ represents the deployment CFP incurred while executing application i on system k. This formulation naturally captures the role of modularity, reconfigurability, and chiplet reuse in amortizing embodied CFP across multiple applications, deployments, and product lifetimes.

The number of deployed system instances, N_{sys}, depends on the target platform's flexibility and reusability. Fixed-function ASIC-based systems may require different hardware systems for evolving workloads or application domains, increasing the number of deployed hardware instances. In contrast, HI systems with FPGA chiplets can support multiple workloads through post-deployment programmability, while also reusing common chiplets, such as memory, networking, connectivity, and accelerator chiplets, across multiple systems and generations. As a result, embodied CFP can be amortized over a broader set of applications, deployments, and lifetimes. In general, N_{sys} varies with the degree of programmability, modularity, and reusability provided by the HI system architecture.

Embodied CFP: Embodied carbon footprint (CFP) captures emissions arising from activities such as material sourcing, chiplet design, wafer fabrication, testing, packaging, integration, assembly, and end-of-life treatment. For HI systems with FPGA chiplets, embodied CFP depends not only on the properties of individual

chiplets, but also on the complexity of integrating a large number of heterogeneous chiplets into a complete system. In particular, technology-node selection impacts fabrication carbon intensity, while chiplet count directly influences package complexity, die-to-die interconnect requirements, assembly overheads, and package-level yield. As the number of chiplets increases, manufacturing CFP may decrease due to improved die-level yield and smaller die sizes, while package-related CFP can increase due to additional bonding steps, substrates, interconnect structures, and assembly complexity.

From a system-level perspective, embodied CFP in HI systems with FPGA chiplets consists of both per-chiplet and package-level contributions. The per-chiplet terms include manufacturing, amortized design, and testing CFP for each chiplet, while the system-level terms capture package fabrication, integration overheads, off-chip memory, and end-of-life processing. Accordingly, the embodied CFP of HI system instance k is given by:

$$C_{\text{emb},k} = \sum_{j=1}^{N_c} \left(C_{\text{mfg},j} + \frac{C_{\text{des},j}}{N_{\text{vol}}} + C_{\text{test},j} \right) + C_{\text{pkg}}(N_c) + C_{\text{mem}} + C_{\text{eol}} \tag{2}$$

where N_c denotes the number of chiplets in HI system instance k, including FPGA, CPU/GPU, memory, networking, connectivity, analog, and accelerator chiplets. $C_{\text{mfg},j}$ is the manufacturing CFP of chiplet j, $C_{\text{des},j}$ is the design CFP of chiplet j amortized over the production volume N_{vol}, and $C_{\text{test},j}$ is the testing CFP of chiplet j. $C_{\text{pkg}}(N_c)$ captures the package and integration-related CFP associated with assembling the N_c chiplets into a complete system, including advanced packaging, die-to-die interconnects, bonding, substrates, and assembly overheads. C_{mem} and C_{eol} denote the contributions of off-chip memory and end-of-life processing, respectively. This formulation captures the key sustainability tradeoff in HI systems: increasing chiplet count can improve yield and reuse, while simultaneously increasing package-related embodied CFP. The detailed models for each of these terms are available in [25, 26].

Deployment CFP: The deployment CFP $C_{\text{deploy},k,i}$ captures the carbon footprint incurred after hardware production during deployment of application i on HI system instance k. In HI systems with FPGA chiplets, deployment CFP consists of both operational CFP during system use and the CFP associated with adapting and configuring the FPGA chiplets and surrounding system components for target workloads. The deployment CFP is given by:

$$C_{\text{deploy},k,i} = C_{\text{op},k,i} + C_{\text{reconfig},k,i} \tag{3}$$

where $C_{\text{op},k,i}$ accounts for the operational CFP due to energy consumed while executing application i on HI system instance k:

$$C_{\text{op},k,i} = C_{\text{src,use}} \times E_{\text{use},k,i} \tag{4}$$

where $C_{\text{src,use}}$ is the carbon intensity of the energy source during deployment and $E_{\text{use},k,i}$ is the total operational energy consumed while executing application i on HI system instance k.

The reconfiguration CFP term $C_{\text{reconfig},k,i}$ captures the carbon overhead associated with adapting the FPGA chiplets and surrounding software stack for application i on system k. This includes compilation, bitstream generation, mapping, runtime tuning, workload scheduling, and other deployment-time adaptation steps. In HI systems with FPGA chiplets, this term becomes particularly important because the same FPGA fabric can be reused across multiple evolving workloads and applications over the system lifetime. As a result, reconfigurability enables embodied CFP to be amortized across multiple deployments while also enabling runtime adaptability for rapidly evolving AI and data-intensive workloads. Detailed models for each of these metrics can be found in [26, 28].

4 Carbon-aware optimization

Existing design-space exploration and optimization frameworks for heterogeneous and chiplet-based systems are not fully carbon-aware. Tools such as HISIM [33] primarily focus on analytical modeling and performance analysis for heterogeneous 2.5D/3D integrated systems, while ECO-CHIP [25] focuses on modeling embodied and operational CFP for chiplet-based platforms. Similarly, ChipletGym [17] explores chiplet design spaces using reinforcement learning (RL) and simulated annealing (SA), but optimization is primarily driven by traditional design objectives rather than lifecycle sustainability metrics. CarbonPATH [27] takes an important step toward lifecycle-aware optimization by jointly considering embodied and operational CFP during design-space exploration for HI systems. In particular, CarbonPATH explores tradeoffs across architecture, packaging, workload mapping, and technology-node assignment while balancing traditional system metrics with sustainability objectives in HI systems with systolic array chiplets only. However, the introduction of FPGA chiplets significantly expands this optimization space. Unlike fixed-function HI systems, HI systems with FPGA chiplets introduce deployment-time programmability, runtime adaptability, compilation and mapping considerations, FPGA fabric sizing, and workload-dependent reconfiguration behavior. As a result, sustainability-aware optimization for HI systems with FPGA chiplets requires additional cross-stack reasoning across compilers, runtime systems, FPGA mapping, and reconfiguration strategies alongside traditional HI-system design parameters.

A carbon-aware optimization framework for HI systems with FPGA chiplets must therefore jointly optimize conventional system metrics together with sustainability objectives. In particular, optimization must account for traditional metrics such as performance, power, area, cost (PPAC), while simultaneously considering embodied and operational CFP across the system lifecycle. To complement conventional PPAC metrics, it is also useful to evaluate how much useful computation a system delivers per unit carbon. For example, higher-performance systems may require larger chiplet counts, more advanced packaging, or higher operational power, increasing both embodied and operational CFP. To better capture this tradeoff, a performance sustainability index (Perf-SI), originally proposed in prior sustainability-focused computing studies [24, 27], can be defined as Perf-SI = $\frac{\text{Performance}}{\text{Total } CO_2\text{eq}}$ where Performance represents the system throughput for a target workload, and CFP includes both embodied and operational carbon contributions.

Fig. 6 illustrates the variation in normalized Perf-SI as the number of chiplets increases for both (a) 3D HI and (b) 2.5D HI systems. In this experiment, each chiplet contains a 128×128 systolic array with 1024 KB SRAM buffers implemented in a 7nm technology node and connected to DDR5 off-chip memory while executing a vision transformer GEMM workload as illustrated in [27]. The figure compares multiple package technologies across both 3D and 2.5D integration styles. Packages with higher die-to-die bandwidth,

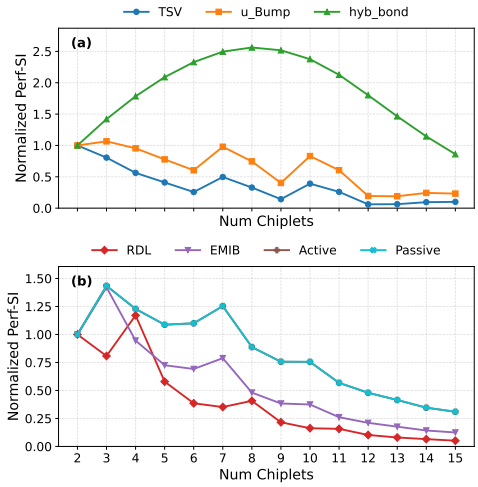


Figure 6: Variation in normalized Perf-SI across different numbers of chiplets, where each chiplet contains a 128×128 systolic array with 1024 KB SRAM in 7nm, for (a) 3D package interconnects and (b) 2.5D package interconnects under a ViT workload. Adapted from [27].

such as hybrid bonding in 3D integration and active/passive interposers in 2.5D integration, achieve larger Perf-SI improvements as chiplet count increases due to reduced communication bottlenecks and higher-bandwidth die-to-die transfers, which improve both performance and operational efficiency. However, as the number of chiplets continues to increase, communication overheads begin to dominate, resulting in reduced overall performance. In such cases, the reduction in CFP achieved through larger disaggregation no longer sufficiently compensates for the performance degradation caused by increased communication and packaging overheads. Although this example uses systolic-array chiplets, it illustrates the type of package- and communication-aware sustainability tradeoff that becomes even richer when FPGA chiplets introduce fabric sizing, reconfiguration, and workload-dependent mapping.

5 Research challenges and opportunities

Several prior works have highlighted broader research opportunities in sustainable computing, sustainability-aware architectures, and environmentally-conscious electronic design automation [2, 4, 5, 22, 29, 36]. This work examines these opportunities from the perspective of HI systems with FPGA chiplets, where modular chiplet-based integration, advanced packaging, and FPGA reconfigurability introduce new sustainability-driven design tradeoffs and optimization challenges across the computing stack.

We identify key research opportunities for sustainable HI systems with FPGA chiplets. These opportunities arise because lifecycle CFP depends not only on architecture and packaging choices, but also on compilation, runtime adaptation, EDA support, manufacturing, reuse, and end-of-life behavior.

Sustainability-aware metrics and modeling: One important research direction is the development of sustainability-aware optimization metrics beyond traditional PPAC objectives. While CFP is an important first step, future HI systems may also require optimization across additional environmental metrics such as water footprint, material criticality, recyclability, rare-earth usage, e-waste generation, and lifecycle reliability [2, 13, 15, 21, 35]. These metrics are often conflicting; for example, reducing operational energy through aggressive technology scaling or advanced packaging may increase embodied CFP due to manufacturing complexity. Such tradeoffs motivate new metrics that account for these conflicts [24].

Architecture and system-level co-design: HI systems with FPGA chiplets create opportunities for sustainability-aware workload specialization and adaptable computing [16, 31]. There are frameworks that support end-to-end framework for automated 3D FPGA architecture generation considering interconnects and packaging [37]. Future systems may dynamically allocate workloads across CPUs, GPUs, FPGA fabrics, memory chiplets, and domain-specific accelerators based on workload behavior, runtime constraints, and sustainability objectives [18]. HI systems also create opportunities for sustainability-aware technology-node assignment [25, 38], where performance-critical digital logic may be fabricated in advanced nodes while analog, I/O, connectivity, and memory chiplets remain in mature technologies. Such heterogeneous integration can reduce unnecessary migration to highly carbon-intensive process nodes while still meeting performance and energy targets.

Chiplet, interconnect, and package co-design: At the package and chiplet-design level, there are significant opportunities for co-design across chiplets, die-to-die interconnects, and package architectures. There are optimization frameworks for mixed-node, mixed-area heterogeneous integration systems that improve PPAC while incorporating thermal awareness across chiplet, interconnect, and package-level design considerations [19]. Future research may explore sustainability-aware chiplet partitioning, topology-aware interconnect optimization, thermal-aware package design, and modular upgrade strategies where portions of a system can be selectively replaced or reused instead of replacing the full platform. The relationship between chiplet granularity, package complexity, yield, thermal behavior, and lifecycle CFP remains a particularly important open challenge.

Compiler, runtime, and reconfiguration optimization: Compiler and runtime-system research becomes increasingly important in HI systems with FPGA chiplets. Compilation, workload mapping, scheduling, FPGA reconfiguration, and runtime adaptation decisions directly impact communication traffic, resource utilization, operational energy, and lifecycle CFP. Emerging programming and compilation frameworks for heterogeneous and reconfigurable systems, such as [14] and [8], further highlight the importance of cross-layer optimization across compilation, HLS, runtime orchestration, and physical design. These approaches can be extended to account for grid carbon intensity, renewable energy availability, cooling demand, and data-center conditions. Similarly, compilers and runtime systems may select FPGA configurations, mapping strategies, or implementation targets based not only on latency and throughput, but also on operational CFP, amortized embodied CFP, workload duration, and expected reuse. These opportunities motivate sustainability-aware compilation and runtime orchestration frameworks for HI systems with FPGA chiplets.

EDA and AI-driven design automation: Existing EDA tools are not designed to optimize lifecycle sustainability metrics. Future

EDA frameworks for HI with FPGA chiplets may need to incorporate embodied and operational CFP directly into optimization across the compute stack. The large design space of HI systems with FPGA chiplets also creates significant opportunities for AI-driven optimization [34]. Traditional heuristic-based approaches may struggle to efficiently explore interacting architecture, packaging, compilation, and runtime decisions involved in sustainable HI-system design. AI-driven frameworks, including machine learning, reinforcement learning [17], foundation models, and agentic optimization systems [34], can potentially enable adaptive design-space exploration while reasoning about sustainability-performance tradeoffs and lifecycle impacts. There are ML-based routablity-aware 3D flow that helps with congestion predictions [11], we also have frameworks that optimize 3D performance and power delivery [40]. Future EDA tools may also support lifecycle-aware hardware design through modular interfaces, repairability, upgradeability, and design-for-disassembly to improve reuse [39].

Manufacturing, reuse, and circular lifecycle management: Sustainable HI systems also create opportunities for research into reusable chiplet ecosystems, circular hardware supply chains, recyclable package materials, remanufacturable systems, and end-of-life-aware hardware design [32, 39]. Since HI systems separate functionality into modular chiplets, future systems may support selective reuse, refurbishment, replacement, or upgrading of individual chiplets, extending system lifetime while reducing lifecycle CFP. FPGA chiplets further strengthen this opportunity by enabling post-deployment adaptability without requiring full hardware redesign. Future sustainability optimization frameworks may therefore treat reuse, refurbishment, upgradeability, and end-of-life recovery as active design objectives rather than fixed terminal costs.

6 Conclusion

HI systems with FPGA chiplets offer a promising pathway toward sustainable computing by combining modular chiplet-based integration with deployment-time reconfigurability. Their sustainability benefits, however, are not automatic; they depend on careful co-optimization of chiplet partitioning, package architecture, FPGA fabric sizing, workload mapping, reuse, and operational behavior across the full lifecycle. Enabling this vision will require new modeling, optimization, compiler, runtime, and EDA methodologies that treat sustainability as a first-class design objective.

References

[1] D. S. Berger et al. 2023. Reducing Embodied Carbon is Important. SIGARCH Blog.
[2] L. Boakes et al. 2023. Cradle-to-gate life cycle assessment of CMOS logic technologies. In *Proc. IEDM*.
[3] A. A. Chien. 2023. Why Embodied Carbon is a Poor Architecture Design Metric, and Operational Carbon Remains an Important Problem. SIGARCH Blog.
[4] P. Dangi et al. 2025. Sustainable Hardware Specialization. In *Proc. ICCAD*.
[5] L. Eeckhout. 2025. Assessing Processor Sustainability Using the First-Order FOCAL Carbon Model. *IEEE Micro* 45, 4 (2025), 11–18.
[6] M. Elgamal et al. 2023. Design space exploration and optimization for carbon-efficient extended reality systems. *arXiv preprint arXiv:2305.01831* (2023).
[7] C. Freitag et al. 2021. The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. *Patterns* 2, 9 (2021).
[8] L. Guo et al. 2023. Tapa: A scalable task-parallel dataflow programming framework for modern FPGAs with co-optimization of HLS and physical design. *ACM Trans. on Reconfig. Tech. and Sys.* 16, 4 (2023), 1–31.
[9] U. Gupta et al. 2022. ACT: Designing Sustainable computer systems with an Architectural Carbon Modeling tool. In *Proc. ISCA*.
[10] E. Haque et al. 2025. EDA for Heterogeneous Integration. In *Proc. DAC*.
[11] H.-H Hsiao et al. 2025. DCO-3D: Differentiable Congestion Optimization in 3D ICs. In *Proc. DAC*.
[12] J. Hu et al. 2025. CarbonSet: A Dataset to Analyze Trends and Benchmark the Sustainability of CPUs and GPUs. In *Proc. GLS-VLSI*.
[13] Y. Jiang et al. 2025. ThirstyFLOPS: Water Footprint Modeling and Analysis Toward Sustainable HPC Systems. In *Proc. ICHPCNSA*.
[14] Y.-H. Lai et al. 2019. HeteroCL: A multi-paradigm programming infrastructure for software-defined reconfigurable computing. In *Proc. ISFPGA*.
[15] P. Li et al. 2025. Making AI less 'Thirsty'. *Commun. ACM* 68, 7 (2025), 54–61.
[16] X. Li. 2022. *High-Performance FPGA-accelerated Chiplet Modeling*. Xingyu-Dissertation-2022.pdf.
[17] K. Mishty and M. Sadi. 2025. Chiplet-Gym: Optimizing Chiplet-Based AI Accelerator Design With Reinforcement Learning. *IEEE T. Computers* (2025), 43–56.
[18] M. Odema et al. 2024. SCAR: Scheduling Multi-Model AI Workloads on Heterogeneous Multi-Chiplet Module Accelerators. In *Proc. MICRO*. 565–579.
[19] M. G. Park et al. 2025. MIX-3D: AI-based Architecture-Circuit Co-design Methodology for Mixed-Node, Mixed-Area 3D ICs. In *Proc. ISLPED*.
[20] D. Patel. 2020. Die Size And Reticle Conundrum – Cost Model With Lithography Scanner Throughput. Semianalysis Newsletter.
[21] J. Roelandts et al. 2025. The Architectural Sustainability Indicator. *IEEE Computer Architecture Letters* 24, 2 (2025), 205–208.
[22] T. Shi et al. 2026. Sustainability of Computing Systems: A Survey from Environmental Impact Perspectives. *TechRxiv* 2026, 0211 (2026). https://www.techrxiv.org/doi/abs/10.36227/techrxiv.177083681.16247479/v1
[23] IEEE Electronics Packaging Society. 2020. Heterogeneous Integration Roadmap. https://eps.ieee.org/technology/heterogeneous-integration-roadmap.
[24] C. C. Sudarshan et al. 2024. Beyond the Surface: The Necessity for Detailed Metrics in Corporate Sustainability Reports. In *Proc. IGSC*. IEEE.
[25] C. C. Sudarshan et al. 2024. ECO-CHIP: Estimation of Carbon Footprint of Chiplet-based Architectures for Sustainable VLSI. In *Proc. HPCA*.
[26] C. C. Sudarshan et al. 2024. GreenFPGA: Evaluating FPGAs as Environmentally Sustainable Computing Solutions. In *Proc. DAC*.
[27] C. C. Sudarshan et al. 2026. CarbonPATH: Carbon-aware pathfinding and architecture optimization for chiplet-based AI systems. *arXiv preprint arXiv:2603.03878* (2026).
[28] C. C. Sudarshan et al. 2026. Evaluating Computing Platforms for Sustainability: A Comparative Analysis of FPGAs against ASICs, GPUs, and CPUs. *arXiv preprint arXiv:2604.20638* (2026).
[29] J. Sun et al. 2026. An Analytical Model for Performance-Carbon Co-Optimization of Edge AI Accelerators. *IEEE T. Comput. Aid. D.* 45, 3 (2026), 1499–1503.
[30] T. Tan. 2021. *System level tradeoffs between ASIC and FPGA accelerators*. The University of Texas at Austin. Tan-Dissertation-2021.pdf.
[31] W. Tang et al. 2024. Arvon: A Heterogeneous System-in-Package Integrating FPGA and DSP Chiplets for Versatile Workload Acceleration. *IEEE J. Solid-St. Circ.* 59, 4 (2024), 1235–1245.
[32] I. Tashdid et al. 2025. ECOLogic: Enabling Circular, Obfuscated, and Adaptive Logic via eFPGA-Augmented SoCs. In *Proc. ICCD*.
[33] Z. Wang et al. 2025. HISIM: Analytical Performance Modeling and Design Space Exploration of 2.5D/3D Integration for AI Computing. *IEEE T. Comput. Aid. D.* (2025), 3208–3221.
[34] Q. Wu et al. 2026. CHICO-Agent: An LLM Agent for the Cross-layer Optimization of 2.5D and 3D Chiplet-based Systems. *arXiv preprint arXiv:2604.18764* (2026).
[35] Y. Wu et al. 2025. Not all water consumption is equal: A water stress weighted metric for sustainable computing. *ACM SIGENERGY Energy Informatics Review* 5, 2 (2025), 84–90.
[36] Z. Yin et al. 2025. Toward Lifelong-Sustainable Electronic-Photonic AI Systems via Extreme Efficiency, Reconfigurability, and Robustness. In *Proc. ICCD*.
[37] I. Youssef et al. 2025. LaZagna: An Open-Source Framework for Flexible 3D FPGA Architectural Exploration. In *Proc. ICCAD*.
[38] Y. Zhao et al. 2024. 3D-carbon: An analytical carbon modeling tool for 3D and 2.5D integrated circuits. In *Proc. DAC*.
[39] P. Zhou et al. 2023. REFRESH FPGAs: Sustainable FPGA chiplet architectures. In *Proc. IGSC*.
[40] L. Zhu et al. 2024. Hetero-3D: Maximizing Performance and Power Delivery Benefits of Heterogeneous 3D ICs. In *Proc. ISLPED*.