Timing and Power Optimization in Chiplet-Based SoCs with Heterogeneous Interconnects

Phaneendra Chainulu Sri Adibhatla

Independent Researcher

Jawaharlal Nehru Technological University Hyderabad, India

***Abstract*— The traditional architecture of System-on-Chip (SoC) has changed with the launch of chiplet-based architectures, enabling modular System-on-Chip integration and providing benefits in scalability, performance, and design flexibility. The review analyzes recent advances and the timing and power optimization problems of chiplet-based SoCs, particularly when heterogeneous interconnects are involved. Because chiplets may operate at different voltages, frequencies, and process nodes, timing synchronization and effective power distribution are essential. The article summarizes recent studies on modular AI acceleration, interconnect protocols, cost–performance trade-offs, and power management frameworks, with special emphasis on heterogeneous integration approaches such as 2.5D and 3D packaging. New tools and protocols, such as ChIP and decentralized power control mechanisms, are mentioned with respect to their role in reducing latency, energy consumption, and design complexity. Moreover, the article discusses developments in design automation and semiconductor technologies that support resilient timing closure and power optimization within multi-chiplet frameworks. The review indicates the significance of standardization, architectural co-design, and predictive EDA tools in realizing high-performance, energy-efficient chiplet-based SoCs suitable for next-generation computing systems.**

***Index Terms*— chiplet-based SoC, timing optimization, power management, heterogeneous interconnects.**

I. INTRODUCTION

Scaling issues have severely curtailed the further extension of Moore's Law, and new architecture-level approaches have been taken in the design and implementation of Systems-on-Chip (SoC) architectures. These innovations include chiplet-based SoCs, which have been proposed as a viable solution by subdividing monolithic chips into smaller, reusable, and heterogeneous modules referred to as chiplets. These chiplets are linked by advanced packaging and interconnect technologies, which makes system integration scalable and efficient. As this evolves, the timing and power optimization of such architectures become important factors of consideration, particularly with the complexity that heterogeneous interconnects and modular architectures introduce.

As the semiconductor industry transforms to the paradigm of chiplet-based architectures as opposed to the monolithic integration models of earlier years, concerns arise regarding how to coordinate timing and manage power across different chiplet models. This is further complicated by the heterogeneity of interconnects with varying latency, bandwidth, and power characteristics. Addressing these issues requires a deep familiarity with architectural features, communication principles, and power-distribution systems that regulate communication between chiplets.

The purpose of this review paper is to provide a critical review of timing and power optimization techniques in chiplet-based SoCs with heterogeneous interconnects. In this paper, we rely on the latest literature to discuss architectural changes, performance–space trade-offs, dynamic power management methods, and interface protocols that define this emerging research area.

2. Evolution of Chiplet-Based SoCs

Chiplet-based architectures rearchitect traditional SoC designs to be modular, reusable, and vertically integrated. Functions are built as chiplets, as components of a larger system, such as CPUs, GPUs, accelerators, and memory controllers, and placed on a package substrate. Advantages of this modular approach include enhanced yield, reduced design complexity, and process technological flexibility. Each chiplet can be produced at the most suitable technology node based on its function, which reduces the overall cost and improves the performance of the system [1].

One of the benefits of chiplet integration is its capability to address the growing design and production costs of new technology nodes. With increased transistor densities, thermal density, power delivery, and interconnect delay challenge monolithic SoCs. Chiplets circumvent this issue by permitting the spatial disaggregation of computing and memory modules, which allows greater thermal circulation and improved routing performance. In addition to this, chiplet architectures enable shorter time-to-market since a tested chiplet can be reused across diverse product and system architectures [1].

These systems further facilitate architectural specialization that allows the incorporation of application-specific chiplets to meet certain computational needs such as machine learning, graphics processing, or security. This heterogeneity contributes to the heterogeneous nature of chiplet SoCs and increases the flexibility of the systems, at the cost of increased design complexity

regarding interconnection and power optimization. The interconnect fabric that connects the chiplets in such systems is the most important determinant of timing and power in the system.

3. Modular Integration and AI Acceleration

The increasing requirement to accelerate AI in edge and datacenter appliances is one of the factors driving the growing use of chiplet-based designs. AI-specific accelerators may be dynamically deployed as separate chiplets in modular SoCs. This modularity not only allows AI workloads to be better customized, but also enables selective power-gating and frequency scaling based on workload needs. The separation of AI engines from the central computation logic permits autonomous voltage and frequency regulation, which enables power optimization [2].

The open-source nature of the instruction set can assist in modular AI acceleration in systems using the RISC-V architecture. Designers have the capability of including custom chiplets that can enhance AI inference and training without proprietary IP cores. The RISC-V SoCs presented in [2] reveal the flexibility of SoC design, where data exchange and timing management among AI chiplets and other modules are simplified by the latency-conscious design of interconnects.

A significant challenge of these designs is managing communication overhead between chiplets, especially when AI workloads require high-throughput data. To achieve data coherence, timing synchronization is necessary to avoid pipeline stalls. Dynamic voltage and frequency scaling (DVFS) and intelligent workload partitioning between chiplets help control power consumption. The modular approach also offers the capability of selectively powering down idle chiplets, which achieves significant energy savings without performance degradation [2].

4. Heterogeneous Chiplet Integration for ML and Edge Devices

The scalability of machine learning (ML) applications at the edge is made possible by heterogeneous integration. Chiplets customized for edge ML workloads are enabled by heterogeneous process technologies, such as the integration of high-performance compute dies with low-power sensor or signal-processing dies. These hetero-chiplet systems offer scalability in configuring computing devices to suit a specific workload while minimizing power consumption [3].

These systems rely on the optimization of timing through fine-grained coordination between chiplets with varying operating rates and latencies. The architectural design in [3] suggests silicon interposers and active bridges as communication media between chiplets in order to ensure timing integrity within the SoC. By supporting variable clock domains and asynchronous interfaces, these interconnection media enable reliable data transfer between chiplets with different voltage and timing requirements.

Power efficiency is achieved with chiplets containing custom power domains, where they can operate in optimal power–performance regions. In addition, the placement of near-memory compute chiplets in proximity to data-rich modules reduces the overhead of data movement, which in turn results in reduced power consumption. The timing and power optimization techniques also utilize adaptive clocking schemes that respond to workload stresses, ensuring predictable latency as well as limited jitter across the system [3].

5. Performance Trade-Offs in Chiplet-Based Systems

Though chiplet designs offer advantages in terms of architecture, performance trade-offs must be considered. Latency between chiplets due to modular integration, power waste because of interconnects between chiplets, and synchronization issues between autonomous chiplets can negate the benefits of modular integration. These issues have been raised in the performance analysis presented in [4], which gives insight into the trade-offs that are allowable between interconnect latency, bandwidth, and power.

Some of the types of packaging the paper evaluates include 2.5D and 3D integration to discover their impact on timing and power measurements. However, 2.5D packaging using silicon interposers has lower latency than organic substrates, whereas 3D stacking provides higher vertical bandwidth, though with higher thermal density and power management complexity. Interconnect protocols also influence timing synchronization and power dissipation among chiplets [4].

The graph in **Figure 1** below shows the comparative performance of different packaging strategies based on system latency and power consumption:

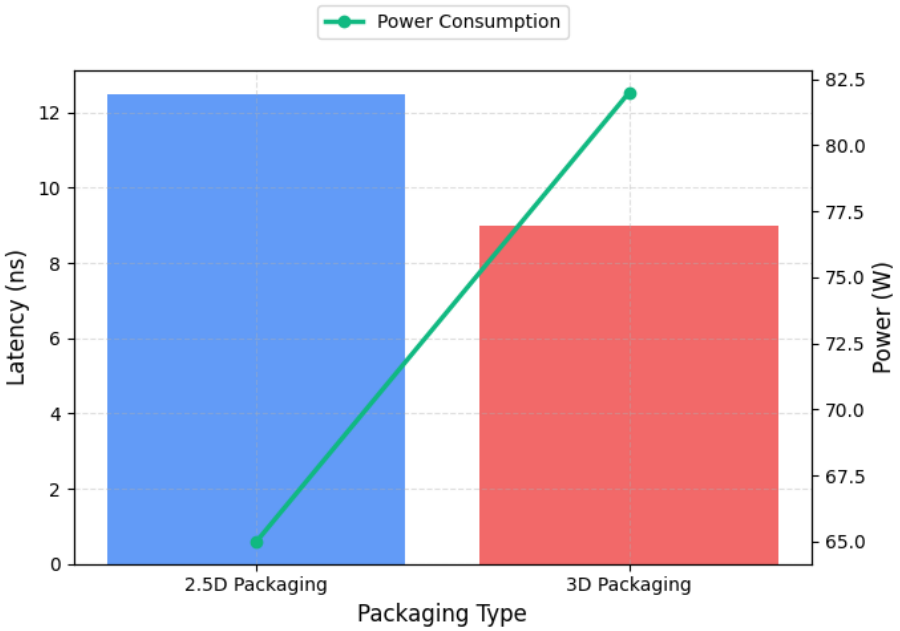

Figure 1: Latency vs Power Consumption for 2.5D and 3D Packaging in Chiplet SoCs
Adapted from [4]

Figure shows trade-offs between latency and power in chiplet-based SoCs using various packaging options.

From the graph, it is evident that while 3D integration offers performance benefits in terms of bandwidth, the associated power costs and thermal implications must be addressed through enhanced cooling techniques and power-aware design methodologies.

6. Cost and Power Co-Optimization Frameworks

Co-optimization of cost, timing, and power in heterogeneous systems is a complex process. The trade-offs are captured in architectural models and simulators such as CATCH (Cost Analysis Tool for Chiplets) used in [5]. CATCH evaluates the system-wide performance and power consumption of a SoC in relation to varying values of several parameters, including chiplet location, interconnect technology, and process node selection.

Both analytic and empirical models are contained in this tool in order to predict power consumption as a function of interconnection topology and chiplet communication patterns. The timing paths are represented at both intra-chiplet and inter-chiplet delays to enable timing closure to be studied at a very early design stage. Power-gating and DVFS models are used in the simulation environment to test the energy efficiency of workloads under different loads [5].

Table 1 below summarizes the optimization parameters and objectives handled by the CATCH framework:

Table 1: Optimization Parameters in CATCH for Chiplet-Based SoCs
Adapted from [5]

Parameter	Optimization Goal	Impact Area
Chiplet Placement	Minimize routing complexity	Power, Timing
Interconnect Type	Reduce transmission delay	Timing, Latency
Process Node Selection	Lower power/performance cost	Power Efficiency
Power Domain Partitioning	Enable DVFS and power gating	Energy Consumption
Packaging Strategy	Balance cost and performance	Cost, Thermal Handling

These tools need to be applied at the early stages of design to prevent the expensive development at the physical design stage. The designers will be able to predict and anticipate the performance constraints and explore alternative integration schemes through simulation of the different settings and operating conditions.

7. Protocols for Chiplet Interconnects

Chiplet-based SoCs are characterized by a collection of components manufactured by various processors and vendor sets; therefore, the standardization of the interconnect protocol is critical to guaranteeing effective data transfer and time synchronization. A new protocol to scale interconnect communication is the Chiplets Interface Protocol (ChIP), which is aimed at ultra-large-scale scenarios [6]. The heterogeneity of chiplet environments is addressed by ChIP, offering a single and scalable means of providing high-speed communication with minimal timing overhead.

ChIP facilitates heterogeneous communication between chiplets with customizable data link layers and robust clock-domain crossing (CDC) mechanisms. These are important for maintaining timing integrity when chiplets operate at varying frequencies. The protocol also includes flow control and error correction mechanisms that reduce the impact of jitter, while offering low-latency communication paths through the interposer or substrate. Handshake synchronization is performed through timing optimization and elastic buffers that reduce the impact of asynchronous communication between chiplets [6].

In addition to its timing advantages, ChIP provides benefits for power optimization. The protocol allows adaptive power states between chiplets with respect to traffic patterns so that idle chiplets or unused lanes can enter low-power states. This type of dynamic power control is critical in multi-die systems, where all modules need not remain active at all times. Moreover, ChIP supports scalable topologies, allowing additional chiplets to be integrated in the future without compromising timing closure or exceeding power budgets [6].

8. 3D Integration and Its Impact on Power and Timing

The use of three-dimensional (3D) SoCs is one strong argument as to why timing and power need to be further optimized in chiplet-based architectures. In 3D SoCs, interconnect delay and power consumption are reduced because of the vertical stacking of chiplets, which leads to shorter data paths. Furthermore, the signal integrity and bandwidth density of 3D integration are better than those of 2.5D or organic interconnect approaches [7].

However, there are also complex timing problems related to 3D integration, particularly in clock distribution and thermal management. Temperature variation across layers can cause clock skew and increased timing uncertainty, especially in high-speed designs. These concerns are addressed by adopting advanced techniques such as Through-Silicon Vias (TSVs), micro-bump interconnections, and central timing controllers in 3D SoCs. These techniques help maintain clock synchronization across stacked layers [7].

In terms of power, the low interconnect power of 3D SoCs has the disadvantage that the high connectivity of active dies can raise thermal hotspots. This generates the need for distributed voltage regulation and active cooling. Voltage islands and fine-grained DVFS are typically used to adaptively scale performance based on thermal conditions. Figure 2 illustrates a typical 3D chiplet-based SoC design that incorporates thermal-aware and timing-optimized layout strategies.

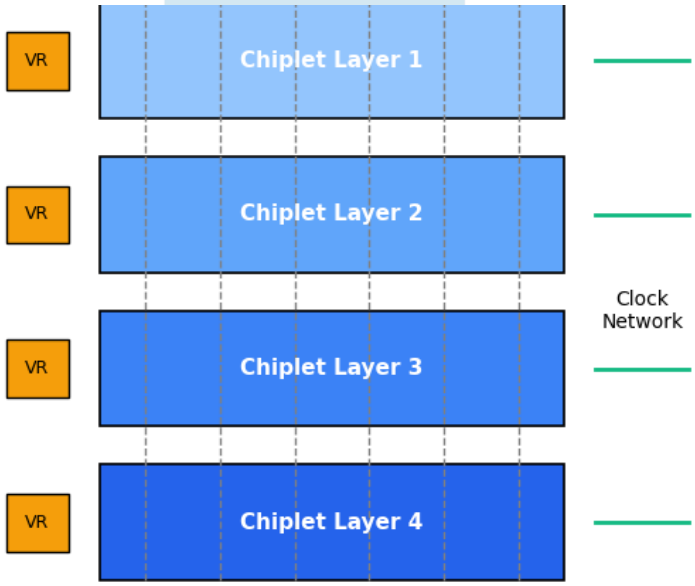

Figure 2: Diagram of a 3D Chiplet-Based SoC with Thermal and Timing Optimization Layers
Adapted from [7]

The diagram shows stacked dies with inter-die vias, embedded voltage regulators, and clock distribution networks for maintaining timing integrity and thermal balance.

The introduction of thermal sensors and power-conscious schedulers in 3D SoCs also enables optimisation of dynamic power. These sensors pass live information to the system controller which in turn readjusts the operation parameters of each chiplet to balance performance and thermal output.

9. Decentralized Power Management Techniques

With the increasing complexity of chiplet-based SoCs, conventional centralized power management schemes are unable to scale. One of the methods proposed in [8] is a decentralized hardware method, which involves the inclusion of distributed power management agents embedded within each of the chiplets. Based on the nature of the workloads, power budgets, and thermal information, these agents generate local decisions and thus enable finer control over the power state of the SoC.

The decentralized model is more receptive to changes in workload, and the overhead associated with centralized control loops is reduced. Power gating, DVFS, and frequency throttling can also be executed locally by the agents at their respective locations. Such architectures also enhance fault tolerance because the failure of power management in one chiplet does not impact the entire SoC [8].

The adoption of blockchain-inspired consensus mechanisms can also be used to allow chiplets to make global power decisions jointly without requiring a single point of control, thereby increasing power efficiency. These lightweight schemes do not require continuous synchronization of the global state, which typically leads to large timing and energy overheads in large-scale systems.

Decentralized systems also have the advantage of reduced latency in power-state changes, which allows the SoC to respond to performance requirements at lower energy costs. This solution supports the migration toward more autonomous SoC architectures that can self-optimize using real-time data without routine software-level interventions [8].

10. Design Automation and Integration Challenges

The development of chiplet-based architectures requires electronic design automation (EDA) tools, particularly to coordinate timing and power optimization of heterogeneous modules. Traditional monolithic design tools are ill-suited to handle the diverse process nodes, asynchronous domains, and thermal gradients of chiplet-based systems. As stated in [9], existing EDA tools should be re-evaluated to support chiplet-aware partitioning, placement, and routing.

Timing closure across different dies, clock synchronization across different clock domains that do not always coincide, and power delivery network (PDN) optimization across dies are some of the significant problems of chiplet integration. These challenges are further complicated by the existence of heterogeneous systems in which chiplets are developed according to varying design standards and protocols. In EDA, therefore, the tools employed must contain advanced models of interconnect parasitics, latency prediction, and multi-physics analysis in order to produce valid timing and power analysis [9].

Chiplet-based SoCs are becoming more scalable with innovations in integration-aware synthesis, interconnect planning, and hierarchical design validation. Automation in floorplanning and package-level design helps optimize the physical proximity of high-traffic chiplets at both die and package scales, reducing interconnect delay and power loss. Cross-domain simulation tools are also useful in identifying potential timing bottlenecks and power hotspots in the early design phase, which reduces post-silicon validation costs [9].

In addition, EDA tools supported by AI are being applied to predict timing violations and power anomalies based on historical design data, significantly speeding up the design loop. The combination of AI and traditional EDA workflows also aids in exploring unconventional chiplet topologies and power optimization configurations that would otherwise be difficult to discover using rule-based design flows.

11. Advances in Semiconductor Technology for Chiplet Optimization

Recent advances in semiconductor technology are now contributing significantly to the success of chiplet-based SoCs, both in the fabrication process and in the integration process. More modern materials, such as high-k dielectrics and advanced interconnect metals, have been used to minimize resistance–capacitance (RC) delays and power dissipation in inter-chiplet communication, as reported in [10].

Additionally, smaller-pitch interconnects and higher-density stacking of chiplets have been enabled by new lithography methods and extreme ultraviolet (EUV) processes. This contributes directly to higher timing and energy efficiency. New packaging technologies, including the embedded multi-die interconnect bridge (EMIB) and fan-out wafer-level packaging (FOWLP), allow high-speed, low-power communication between chiplets without negatively impacting signal integrity [10].

The change in the industry toward the utilization of chiplet standards such as UCIe (Universal Chiplet Interconnect Express) is also affecting design methodologies. These standards have helped achieve interoperability and have simplified design complexity, allowing manufacturers to focus more on improved power and timing performance rather than compatibility. In addition, heterogeneous integration platforms allow the combination of logic, memory, RF, and photonic chiplets within the same SoC, extending system capability within power budgets [10].

All of these developments can be employed to achieve the goal of timing closure and energy-efficient execution in sophisticated, modular SoCs. Process technologies and design tools are important in breaking the constraints of traditional scaling strategies and meeting the needs of next-generation applications.

12. Conclusion

The adoption of chiplet-based SoCs represents a paradigm shift in semiconductor design, as scalability, versatility, and performance efficiency necessitate their use. Timing and power optimization in such architectures is a complex but very important goal, particularly when heterogeneous interconnects are involved. This review acknowledges advances in architecture, communication protocols, power management, and toolchain developments that enable efficient integration of systems using chiplets.

Enabling factors such as ChIP protocols, decentralized power agents, 3D integration, and advanced packaging processes are essential in addressing synchronization and energy management challenges. The ability to model, simulate, and optimize chiplet-based SoCs is further enhanced by the integration of EDA tools optimized for heterogeneous environments. Continued progress in semiconductor manufacturing and standardization efforts will further drive the adoption of chiplet technologies and establish a strong platform for future computing systems that are not only high-performing but also energy-efficient.

References

[1] Mohammed, H. U. H. (2025). CHIPLET-BASED ARCHITECTURES: REDEFINING THE FUTURE OF SYSTEM-ON-CHIP (SOC) DESIGN.
[2] Ramkumar, P., & Bharadwaj, S. S. (2025). Chiplet-Based RISC-V SoC with Modular AI Acceleration. *arXiv preprint arXiv:2509.18355*.
[3] Adiletta, M. J., Wei, G. Y., & Brooks, D. (2025). Democratizing Customization for ML at the Edge through Hetero-Chiplet SiP Architectures. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*.
[4] Mallya, N. B., Strikos, P., Goel, B., Ejaz, A., & Sourdis, I. (2025, March). A Performance Analysis of Chiplet-Based Systems. In *2025 Design, Automation & Test in Europe Conference (DATE)* (pp. 1-7). IEEE.
[5] Graening, A., Talukdar, J., Pal, S., Chakrabarty, K., & Gupta, P. (2025). CATCH: a Cost Analysis Tool for Co-optimization of chiplet-based Heterogeneous systems. *arXiv preprint arXiv:2503.15753*.
[6] Delavari, A., & Vaisband, B. (2025). Chiplets Interface Protocol (ChIP) for Ultra-Large-Scale Applications. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*.
[7] Chakravarthi, V. S., & Koteshwar, S. R. (2025). Designing Three-Dimensional System on Chips (3D SOCs). In *SOC-Based Solutions in Emerging Application Domains* (pp. 39-49). Cham: Springer Nature Switzerland.
[8] Cochet, M., Swaminathan, K., Loscalzo, E., Zuckerman, J., dos Santos, M. C., Giri, D., ... & Bose, P. (2025). BlitzCoin: A Decentralized Hardware Solution for Power Management of Highly-Heterogeneous SoCs. *IEEE Micro*.
[9] Haque, E., Nalla, P., Sudarshan, C. C., Yogi, D., Zhang, H., Chakrabarti, C., ... & Sapatnekar, S. S. (2025, June). EDA for Heterogeneous Integration. In *2025 62nd ACM/IEEE Design Automation Conference (DAC)* (pp. 1-4). IEEE.
[10] Keene, B. (2025). Recent Advances in Semiconductor and Microprocessor Technologies. *Transactions on Computational and Scientific Methods*, *5*(8).