A Systematic Literature Review of Network-on-Chip Architectures: Topologies, Routing, Reliability, and Emerging Technologies

Asrani Lit(✉), Rohana Sapawi, Shamsiah Suhaili, Sharifah Masniah Wan Masra, and Kuryati Kipli

Department of Electrical Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Malaysia
{lasrani,srohana,sushamsiah,wmmasnia,kkuryati}@unimas.my

Abstract. The increasing core count of contemporary systems-on-chip (SoCs) has driven a fundamental shift from shared bus architectures to scalable on-chip interconnection networks. Network-on-Chip (NoC) has emerged as the de facto paradigm for providing high-bandwidth, low-latency, and energy-efficient communication among heterogeneous processing elements. Over two decades of research have produced a rich body of work on NoC topologies, routing algorithms, flow-control schemes, quality-of-service (QoS) mechanisms, reliability enhancement, and emerging technologies such as wireless, optical, and three-dimensional (3D) NoCs. This paper presents a systematic literature review (SLR) of Network-on-Chip architectures from 2002 to 2025 following a transparent and reproducible methodology. We formulate five research questions covering research themes, design challenges, technological trends, evaluation methodologies, and open problems. A structured search over major digital libraries is conducted, and primary studies are classified along eight dimensions: topology design, routing algorithms, flow control and buffering, QoS and real-time guarantees, fault tolerance and reliability, energy and thermal awareness, heterogeneous and emerging NoCs, and machine-learning-driven NoC optimization. The review highlights a progression from regular 2D meshes with deterministic routing to heterogeneous hybrid interconnects with adaptive, often learning-based routing and cross-layer co-optimization. We identify research gaps in holistic security, standardized benchmarks for AI-centric workloads, integrated ML frameworks for NoC design space exploration, and cross-technology co-design in hybrid wired–wireless–optical systems. The findings provide a consolidated reference for researchers and practitioners, and outline promising directions for next-generation NoC research.

A. Lit, R. Sapawi, S. Suhaili, S. M. Wan Masra and K. Kipli—Contributed equally to this work.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026, LNCS 16767, pp. 375–390, 2027.
https://doi.org/10.1007/978-3-032-30491-9_24

Keywords: Network-on-Chip (NoC) · routing · topology · reliability · wireless NoC · optical NoC · 3D NoC · systematic literature review (SLR) · machine learning

1 Introduction

The transition to manycore and heterogeneous system-on-chip (SoC) platforms has imposed stringent demands on the on-chip interconnection fabric. Shared buses and crossbars, once adequate for a small number of cores, fail to scale gracefully due to bandwidth limitations, complex arbitration, and prohibitive power and area overheads [1]. Network-on-Chip (NoC) emerged as a scalable and modular on-chip communication paradigm by leveraging concepts from off-chip interconnection networks and packet-switched communication [2].

NoC-based interconnects have been widely explored in academic research and adopted in industrial designs, particularly in multicore CPUs, GPUs, and heterogeneous SoCs. Research has spanned a broad spectrum, including topology design [3,4], routing algorithms [5,6], flow control and buffering [7,8], quality-of-service (QoS) mechanisms [9], reliability and fault tolerance [10,11], energy and thermal management [12,13], and emerging technologies such as wireless NoC (WiNoC) [14], optical NoC (ONoC) [15], and 3D NoC [16,17]. More recently, machine learning (ML) has been applied to NoC design space exploration and adaptive control [18,19].

Despite the breadth of work, the literature remains fragmented across subtopics and technologies. Several surveys exist [20–22], but many predate the latest developments in hybrid wired–wireless–optical interconnects, ML-assisted optimization, and AI-centric workloads. Furthermore, a comprehensive and systematically conducted review, following established SLR guidelines, is still lacking for Network-on-Chip.

This paper addresses this gap by providing a systematic literature review of NoC architectures and related design methodologies over the period 2002–2024. We follow a structured process, including research question formulation, search strategy definition, study selection, quality assessment, data extraction, and thematic synthesis.

1.1 Research Questions

These research questions are formulated to provide a structured and comprehensive investigation of the existing body of literature. They aim to identify prevailing research trends, dominant methodologies, and technological advancements reported in prior studies, while also uncovering existing limitations and unresolved challenges. Collectively, the RQs serve as the foundation for the review protocol, ensuring a transparent, reproducible, and objective synthesis of evidence aligned with the objectives of this study. The systematic literature review is guided by the following research questions (RQs):

- **RQ1:** What are the dominant research themes and trends in Network-on-Chip architectures?
- **RQ2:** Which NoC design challenges (latency, throughput, energy, reliability, scalability, QoS) have been most extensively studied?
- **RQ3:** How have emerging NoC technologies (wireless, optical, 3D, hybrid) evolved in the literature?
- **RQ4:** What methodologies, tools, and benchmarks are used to evaluate NoC proposals?
- **RQ5:** What research gaps and future directions can be identified from the existing body of work?

1.2 Contributions

The main contributions of this paper are:

- A systematic and reproducible survey protocol for NoC literature aligned with general SLR guidelines.
- A taxonomy of NoC research organized into eight dimensions: topology, routing, flow control, QoS, reliability, energy/thermal, emerging technologies, and ML-driven optimization.
- Comparative analysis of prominent topologies, routing algorithms, and technology platforms, highlighting trade-offs among area, latency, bandwidth, and energy.
- A synthesis of evaluation methodologies, including simulation frameworks (e.g., BookSim, Noxim, Garnet), analytical models, and experimental platforms.
- Identification of research gaps and promising directions for next-generation NoC systems, including security, AI-driven workloads, and hybrid interconnects.

1.3 Paper Organization

The remainder of this paper is organized as follows. Section 2 presents background on NoC architectures. Section 3 details the SLR methodology. Section 4 presents the classification and thematic synthesis of the selected studies. Section 5 discusses key findings and research gaps. Section 6 outlines future research directions, and Sect. 7 concludes the paper.

2 Background on Network-on-Chip

2.1 Basic NoC Architecture

A typical NoC consists of routers interconnected via links in a chosen topology. Processing elements (PEs), such as CPU cores, accelerators, or memory controllers, connect to the NoC through network interfaces. Packets are segmented

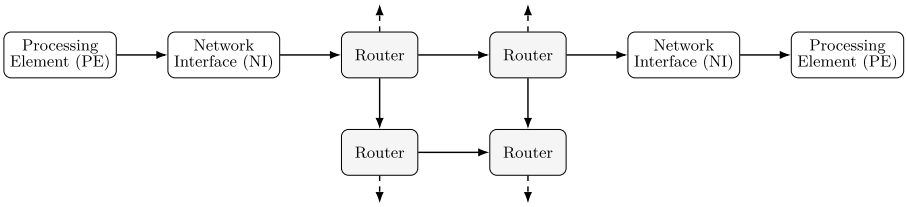

Fig. 1. Basic Network-on-Chip (NoC) architecture showing processing elements (PEs) connected via network interfaces (NIs) to a router-based NoC fabric.

into flits and forwarded across routers using a specific flow-control scheme, often wormhole switching [1]. Routing algorithms determine the path that packets take from sources to destinations, subject to constraints such as deadlock freedom, minimal path length, and congestion avoidance [5].

Figure 1 shows a basic Network-on-Chip architecture comprises processing elements interconnected through network interfaces, routers, and physical links arranged in a regular topology such as a two-dimensional mesh. Data communication is packet-based, with routers performing routing, buffering, and arbitration functions to ensure efficient and scalable on-chip communication. By leveraging networking principles, NoCs overcome the scalability and performance limitations of traditional bus-based interconnects and have become a foundational communication fabric in modern many-core systems.

2.2 Design Objectives and Constraints

As the number of processing elements in modern systems-on-chip continues to increase, efficient on-chip communication has become a critical design challenge. Traditional bus-based interconnects suffer from limited scalability, high contention, and excessive power consumption, motivating the adoption of Network-on-Chip architectures as a structured and scalable communication fabric. Key NoC design objectives include:

- **Performance:** Low average latency, high throughput under realistic workloads.
- **Energy Efficiency:** Minimization of dynamic and static power consumed by routers and links.
- **Scalability:** Support for increasing numbers of cores without disproportionate cost.
- **Reliability:** Robustness against transient and permanent faults.
- **QoS and Real-Time Guarantees:** Predictable performance for latency-critical flows.
- **Technology Constraints:** Impact of process scaling, 3D stacking, and novel interconnect technologies.

These objectives are often conflicting, leading to complex trade-offs that drive research into novel topologies, adaptive routing, power management, and cross-layer optimization.

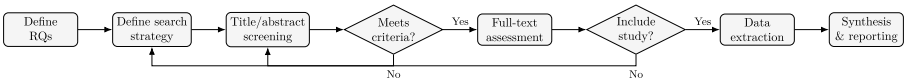

Fig. 2. Systematic literature review (SLR) protocol followed in this work.

3 Systematic Review Methodology

3.1 Protocol and Guidelines

The SLR follows general guidelines such as those by Kitchenham for software engineering SLRs, adapted to hardware and NoC research. The protocol includes the following stages: (i) planning, (ii) conducting, and (iii) reporting. Figure 2 illustrates the overall process followed in this work.

3.2 Search Strategy

Digital libraries searched include IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, and Scopus. The primary search string was constructed using combinations of keywords:

- "network-on-chip" OR "NoC"
- AND ("routing" OR "topology" OR "flow control" OR "QoS" OR "reliability" OR "fault tolerance" OR "wireless" OR "optical" OR "3D")

The search period was set from 2002, corresponding roughly to the emergence of NoC as a research area [2], to the end of 2025. Additional references were obtained via backward and forward snowballing.

3.3 Inclusion and Exclusion Criteria

To ensure the relevance, quality, and consistency of the selected studies, clear inclusion and exclusion criteria were defined prior to the screening process.

Inclusion criteria:

- Peer-reviewed journal and conference papers.
- Primary studies presenting new NoC architectures, algorithms, or analysis.
- Papers focusing on topology, routing, flow control, QoS, reliability, energy/thermal, or emerging NoC technologies.

Exclusion criteria:

- Non-peer-reviewed reports, theses, and patents.
- Works focusing solely on off-chip interconnects or FPGA interconnects without NoC context.
- Papers with insufficient methodological detail or without quantitative or formal analysis.

3.4 Quality Assessment and Data Extraction

Each candidate paper was evaluated using a quality checklist addressing clarity of objectives, soundness of methodology, adequacy of evaluation, and completeness of reporting. For each included study, we extracted:

- Bibliographic data (authors, year, venue).
- NoC dimension(s) addressed (topology, routing, etc.).
- Target technology (2D, 3D, wireless, optical, hybrid).
- Methodology (simulation, analytical modeling, emulation, prototype).
- Evaluation metrics (latency, throughput, energy, area, reliability, QoS).

3.5 Classification Scheme

To systematically analyze and compare the selected studies, the extracted literature was classified into eight primary research categories based on their main technical focus and contribution.

1. **Topology design:** Studies addressing the structural organization of NoC nodes and links, including regular and irregular topologies, with the aim of improving scalability, latency, bandwidth, and physical layout efficiency.
2. **Routing algorithms:** Works proposing deterministic, adaptive, or hybrid routing strategies to optimize path selection, reduce congestion, and ensure deadlock- and livelock-free communication.
3. **Flow control and buffering:** Research focusing on packet switching techniques, buffer management, virtual channels, and congestion control mechanisms to enhance throughput and resource utilization.
4. **QoS and real-time support:** Studies that incorporate quality-of-service guarantees, traffic prioritization, and real-time constraints to support heterogeneous and latency-sensitive applications.
5. **Fault tolerance and reliability:** Contributions addressing transient and permanent faults through error detection, recovery mechanisms, redundant routing, and reliability-aware architectures.
6. **Energy- and thermal-aware design:** Works aiming to minimize power consumption and thermal hotspots via energy-efficient routing, voltage/frequency scaling, and thermal-aware resource management.
7. **Emerging technologies (wireless, optical, 3D, hybrid):** Studies exploring non-traditional NoC implementations, including wireless, photonic, three-dimensional, and hybrid interconnects to overcome the limitations of conventional electrical NoCs.
8. **Machine-learning-driven NoC optimization:** Research leveraging machine learning techniques for adaptive routing, congestion prediction, fault management, and design-space exploration.

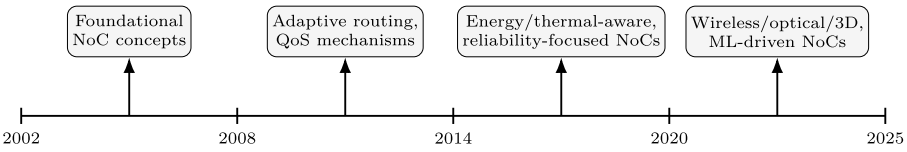

Fig. 3. Conceptual timeline of major research waves in Network-on-Chip (NoC) literature.

4 Results and Thematic Synthesis

4.1 Publication Trends

Figure 3 provides a high-level timeline of key developments in NoC research. Early work focused on regular 2D meshes and deterministic routing [2,20]; the subsequent period introduced topology variants, adaptive routing, and QoS mechanisms [5,9]. The last decade has seen rapid growth in emerging technologies (WiNoC, ONoC, 3D NoC) and ML-assisted design [14,15,18].

4.2 Taxonomy of NoC Research

To provide a structured overview of the research landscape, this study organizes the existing Network-on-Chip literature into a comprehensive taxonomy. The taxonomy captures the major design dimensions, architectural considerations, and optimization objectives that have emerged over time in NoC research. Figure 4 presents the taxonomy used in this review, illustrating the hierarchical classification of studies across core architectural components, communication mechanisms, system-level requirements, and emerging technologies. This taxonomy serves as a unifying framework for analyzing trends, identifying research gaps, and comparing representative works across different NoC domains.

4.3 Topologies

Table 1 summarizes representative NoC topologies and their key characteristics. Mesh remains the most prevalent due to its layout regularity and implementation simplicity [2,20]. Variants such as torus, flattened butterfly, and fat-tree offer reduced hop count at the cost of increased wiring complexity [4]. 3D topologies exploit through-silicon vias (TSVs) to reduce communication distance [16,17].

4.4 Routing Algorithms

Routing algorithms significantly influence NoC performance and robustness. We group routing schemes into deterministic, partially adaptive, fully adaptive, and fault-tolerant categories. Table 2 summarizes representative routing families.

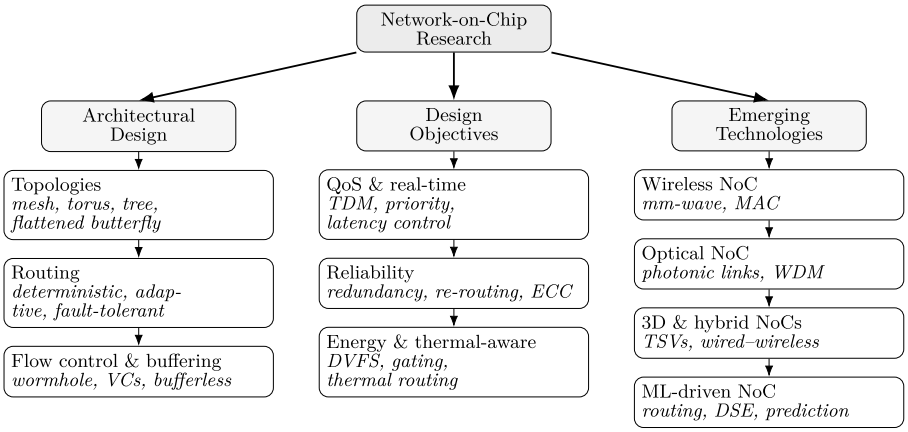

Fig. 4. Taxonomy of Network-on-Chip research.

Table 1. Representative NoC topologies and qualitative comparison.

Topology	Examples	Average hop count	Layout regularity	Routing complexity	Typical applications
2D Mesh	[2,20]	Medium	High	Low (e.g., XY)	General-purpose manycore SoCs
2D Torus	[5]	Low–medium	Medium	Low–medium	Latency-aware systems
Fat-tree / Tree	[3,23]	Low	Medium–low	Medium–high	Hierarchical, clustered SoCs
Flattened Butterfly	[4]	Low	Low	Medium–high	High-radix switches, HPC-style SoCs
Custom / Irregular	[24]	Application-dependent	Low	High	Application-specific NoCs (ASNoCs)
3D Mesh / Torus	[16,17]	Low	Medium	Medium	3D-stacked manycore systems
Hybrid Wired–Wireless	[14,25]	Very low (for long-range flows)	Medium	High	Latency-critical, large-scale NoCs

Table 2. Taxonomy of NoC routing algorithms.

Category	Representative algorithms	Key properties	Typical benefits / limitations
Deterministic	XY, YX, dimension-order [5]	Fixed path for each source–destination pair	Simple hardware, predictable latency; limited congestion handling
Turn-model based	West-first, Negative-first, Odd–Even [26,27]	Restricted turns to avoid cycles	Deadlock-free with modest adaptivity
Partially adaptive	DyAD, minimal adaptive [6]	Adaptive within constraints	Better congestion handling; slightly more complex hardware
Fully adaptive	RCA, OFAR, local congestion-aware [12,28]	High path flexibility	Best performance under load; requires VCs and careful deadlock handling
Fault-tolerant	FTR, LBDR, reconfiguration-based [11,29]	Avoids faulty links/nodes	Improved reliability; additional area and complexity
Wireless/optical-aware	WiNoC-aware and ONoC-aware routing [14,15]	Exploit long-range fast links	Reduced hop count and latency; requires technology-specific design
ML-driven	Reinforcement learning, predictive routing [18,19]	Use ML models to select paths or predict congestion	Potentially higher performance; overhead of training and hardware support

4.5 Flow Control and Buffering

Most NoCs adopt wormhole flow control with virtual channels (VCs) to balance buffer usage and performance [1,8]. Bufferless (deflection) routing is explored to reduce leakage power in deeply scaled technologies [30]. Studies investigate VC allocation, router micro-architecture, and credit-based or on/off flow control mechanisms to improve throughput and energy efficiency.

4.6 QoS and Real-Time Support

QoS mechanisms provide differentiated services for traffic classes with distinct latency, jitter, and bandwidth requirements. Approaches include TDM-based arbitration, priority-based schedulers, and rate control [9,20]. Real-time NoCs for automotive and avionics systems extend these mechanisms with worst-case latency analysis and formal verification [31].

4.7 Fault Tolerance and Reliability

As feature sizes shrink, NoCs become more susceptible to soft errors, wear-out, and TSV failures in 3D stacks. Fault-tolerant techniques include redundancy, fault-aware routing, reconfigurable topologies, and error control coding [10,11,29]. Many works integrate detection and recovery mechanisms to maintain connectivity in the presence of multiple simultaneous faults.

4.8 Energy and Thermal-Aware NoC

Energy consumption arises from dynamic switching in routers and links as well as static leakage. Thermal hotspots degrade reliability and performance. Works in this category propose DVFS, power-gating, clock-gating, and thermal-aware routing [12,13]. In 3D NoCs, vertical communication and hotspot mitigation are central issues [17].

Table 3. Comparison of emerging NoC technologies.

Technology	Key advantages	Main challenges	Typical use-cases	Representative works
Wireless NoC (WiNoC)	Long-range shortcuts, reduced hop count, global broadcast	Interference, antenna design, MAC protocol design, energy of transceivers	Large-scale NoCs, latency-critical traffic	[14,25]
Optical NoC (ONoC)	Very high bandwidth, low latency over long distances	Laser power, thermal sensitivity, fabrication complexity	High-performance computing, bandwidth-intensive SoCs	[15,32]
3D NoC	Shorter distances, reduced footprint, improved performance	TSV reliability, thermal management, manufacturing cost	3D stacked processor-memory systems	[16,17]
ML-driven NoC	Adaptive optimization of routing, power, and resources	Training data requirements, hardware cost of ML models	Dynamic and heterogeneous workloads, AI accelerators	[18,19,33]

4.9 Emerging Technologies and ML-Driven NoCs

Wireless NoC (WiNoC) uses on-chip mm-wave transceivers to create long-range shortcuts that reduce hop count and average latency [14,25]. Optical NoC (ONoC) leverages silicon photonics and wavelength-division multiplexing (WDM) to provide high-bandwidth, low-latency links [15,32]. 3D NoCs exploit TSVs to stack multiple logic and memory layers [16]. ML-driven NoC work uses supervised and reinforcement learning for traffic prediction, adaptive routing, admission control, and design space exploration [18,19,33]. Table 3 summarizes key characteristics of these emerging platforms.

5 Discussion

5.1 Shift from Regular to Hybrid Topologies

While the 2D mesh topology remains dominant in both academic prototypes and commercial manycore processors due to its layout regularity and ease of implementation, the literature reveals a gradual shift toward more complex and hybrid NoC topologies. These include architectures augmented with express links, hierarchical hubs, wireless shortcuts, 3D stacked vertical links, or photonic interconnects to reduce average hop count and communication latency [34,35]. Such hybrid designs are particularly attractive for large-scale and heterogeneous systems, where communication locality and traffic imbalance are pronounced [36,37].

Despite these advantages, industrial adoption still favors regular structures because of their predictable timing behavior, simpler physical design, and lower verification overhead. As a result, hybrid topologies are often explored in constrained or partially deployed forms, highlighting an ongoing trade-off between performance scalability and design complexity.

5.2 From Deterministic to Learning-Based Routing

Routing strategies in NoCs have evolved from deterministic dimension-order schemes, such as XY routing, toward adaptive and congestion-aware algorithms capable of improving throughput and fault tolerance under non-uniform traffic conditions [5]. More recently, machine learning (ML)-based routing approaches have gained increasing attention, replacing hand-crafted heuristics with data-driven policies that learn from runtime traffic patterns and network states [18,19,38].

These learning-based methods, often relying on reinforcement learning or deep neural networks, have demonstrated promising gains in latency reduction and load balancing. However, several challenges hinder their widespread adoption, including hardware overhead, training stability, explainability of routing decisions, and formal verification of correctness and deadlock freedom [39,40]. Addressing these concerns remains a critical research direction for practical deployment.

5.3 Growing Focus on Reliability, Energy, and Thermal Issues

With continued technology scaling and increased integration density, reliability, energy efficiency, and thermal management have emerged as first-order design constraints in NoC architectures. Phenomena such as process variation, aging, soft errors, and thermal hotspots can significantly degrade network performance and lifetime [41,42]. Consequently, recent studies increasingly adopt cross-layer approaches that jointly consider micro-architectural techniques, routing strategies, and system-level management policies [43,44].

This trend reflects a broader shift away from isolated optimizations toward holistic NoC design frameworks that balance performance, power, and reliability objectives simultaneously. Such approaches are especially relevant for 3D-integrated and heterogeneous systems, where thermal coupling and reliability concerns are exacerbated.

5.4 Evaluation Methodologies and Benchmarks

Simulation-based evaluation remains the dominant methodology for NoC research, with widely used frameworks including BookSim, Garnet (Gem5), and Noxim. Most studies rely on synthetic traffic patterns such as uniform random, transpose, or hotspot traffic to stress-test network behavior. A smaller subset employs application-driven traces derived from benchmark suites such as PARSEC, SPLASH-2, or custom workloads.

However, there is still a notable lack of standardized evaluation suites specifically targeting modern AI workloads and heterogeneous accelerator-rich systems. As NoCs increasingly serve as the communication backbone for DNN accelerators and dataflow architectures, recent works have emphasized the need for benchmark traces that capture multicast, bursty, and synchronization-heavy traffic patterns characteristic of AI applications [45–47]. The absence of such standardized benchmarks continues to limit fair comparison and reproducibility across studies.

6 Future Research Directions

The systematic analysis suggests several promising avenues for future work:

6.1 Security and Trust in NoC

Security has only recently begun to receive systematic attention in NoC research, despite the growing deployment of manycore systems in safety- and mission-critical applications. Potential threats include side-channel attacks, denial-of-service via congestion, malicious routers, and hardware Trojans embedded at design or fabrication stages [48,49]. Future work requires robust, low-overhead mechanisms for intrusion detection, secure routing, authentication, and traffic obfuscation that can be integrated without significantly degrading latency or energy efficiency [50,51].

6.2 Standardized AI-Centric Benchmarks

With the proliferation of AI accelerators and dataflow architectures, NoC workloads are increasingly dominated by regular, bursty, and multicast traffic patterns arising from deep neural network (DNN) computations [52,53]. However, many existing evaluations still rely on synthetic or legacy benchmarks that do not accurately capture these characteristics. Community-agreed benchmarks, traffic traces, and evaluation methodologies for AI-centric NoCs would enable more realistic, reproducible, and comparable performance assessments across studies [45,46].

6.3 Integrated ML Frameworks for NoC

Machine learning techniques have been widely explored for adaptive routing, congestion prediction, dynamic voltage and frequency scaling (DVFS), and fault management in NoCs [38,54]. However, these approaches are often applied in an ad hoc manner, focusing on isolated optimization objectives. Integrated ML frameworks that support end-to-end training, deployment, and hardware mapping of ML models for NoC control would accelerate practical adoption. In particular, the co-design of ML algorithms and lightweight hardware accelerators for in-router inference remains an open research challenge [39,55].

6.4 Cross-Technology and Cross-Layer Co-Design

Emerging hybrid wired–wireless–optical NoCs in advanced 2.5D and 3D packages introduce a complex multi-physics design space involving signal integrity, thermal coupling, reliability, and manufacturing cost [34,35]. Addressing these challenges requires holistic co-optimization across circuit, architecture, and application layers. Future research should emphasize cross-layer design methodologies, formal modeling techniques, and scalable design space exploration tools capable of managing this growing complexity [36,56].

6.5 Sustainability and Lifetime Management

Energy-efficient NoC design plays a critical role in improving overall system sustainability and reducing operational carbon footprint [43]. Beyond energy, lifetime-aware routing, wear-leveling, and reliability management techniques can extend the usable lifetime of links, routers, and through-silicon vias (TSVs), particularly in deeply scaled and 3D-integrated technologies [41,42]. Such approaches contribute to reduced electronic waste and improved long-term system cost-effectiveness.

7 Conclusion

This paper has presented a systematic literature review of Network-on-Chip architectures spanning more than two decades of research. Using a structured

protocol, we synthesized contributions across topology design, routing algorithms, flow-control mechanisms, QoS support, reliability and fault tolerance, energy- and thermal-aware techniques, emerging NoC technologies, and ML-driven NoC optimization. We provided taxonomies, comparative tables, and conceptual timelines that highlight the evolution of the field. The identified research gaps and future directions particularly in security, AI-centric benchmarking, ML integration, and cross-technology co-design offer a roadmap for the next generation of NoC research and practice.

Acknowledgement. The authors gratefully acknowledge the support received from the FRGS-EC/1/2024/TK07/UNIMAS/03/1 Malaysia under grant number NAT/F02/FRGS-EC/86558/2024. Additionally, we would like to thank the Faculty of Engineering at Universiti Malaysia Sarawak for their support with the Antenna Laboratory facilities.

References

1. Dally, W.J., Towles, B.: Route packets, not wires: on-chip interconnection networks. In: Proceedings of the Design Automation Conference, pp. 684–689 (2001)
2. Benini, L., De Micheli, G.: Networks on chips: a new SoC paradigm. IEEE Comput. **35**(1), 70–78 (2002)
3. Ni, L.M., McKinley, P.K.: The architecture of interconnection networks. IEEE Comput. **26**(2), 62–76 (1993)
4. Kim, J., Dally, W.J., Scott, S., Abts, D.: Flattened butterfly: a cost-efficient topology for high-radix networks. In: International Symposium on Computer Architecture, pp. 126–137. ACM (2007)
5. Duato, J., Yalamanchili, S., Ni, L.: Interconnection networks: an engineering approach, morgan kaufmann (1997)
6. Puente, V.: Dynamic routing using virtual channels in NOCS. In: International Conference on High Performance Computing (2002)
7. Mutlu, O., Stark, J., Patt, Y.N.: Runahead execution: an alternative to very large instruction windows for out-of-order processors. In: High-Performance Computer Architecture, pp. 129–140. IEEE (2003)
8. Nicopoulos, D., Park, J., Kim, J., Vijaykrishnan, N., Irwin, M., Das, C.: Vichar: A dynamic virtual channel regulation scheme for network-on-chip routers. In: International Symposium on Microarchitecture, pp. 333–346 (2006)
9. Goossens, K., Dielissen, J., Radulescu, A.: Aethereal network on chip: concepts, architectures, and implementations. IEEE Design Test Comput. **22**(5), 414–421 (2005)
10. Srinivasan, S., Gupta, R.: Ras: a reliability-aware scheduler for multiprocessor systems-on-chip. In: Design Automation Conference (2004)
11. Satpathy, A.: FTR: a fault-tolerant router for network-on-chip. In: International Conference on VLSI Design, pp. 243–248. IEEE (2012)
12. Hu, J., Marculescu, R.: Design and exploration of on-chip crossbar networks. In: International Conference on Computer Design, pp. 524–529 (2005)
13. Mercier, P., et al.: Power-aware on-chip interconnect design. IEEE J. Emerg. Select. Topics Circuits Syst. **2**(2), 228–239 (2012)

14. Dang, Q., et al.: A power-efficient wireless network-on-chip architecture. In: International Symposium on Networks-on-Chip, pp. 25–32. IEEE (2011)
15. Pan, D., et al.: Firefly: illuminating future network-on-chip with nanophotonics. In: Design Automation Conference, pp. 694–699. ACM (2010)
16. Patti, R.: Three-dimensional integrated circuits and the future of system-on-chip designs. In: Proceedings of the IEEE, vol. 94, pp. 1214–1224. (2006)
17. Nicopoulos, D.: Data-centric thermal management in 3d multicore architectures. IEEE Trans. Very Large Scale Integr. VLSI Syst. **18**(10), 1432–1445 (2010)
18. Rahmani, A.-M., et al.: Applying machine learning for NOC design space exploration. In: Design, Automation and Test in Europe Conference, pp. 223–228 (2013)
19. Dai, Y.: A convolutional neural network approach for NOC routing. In: International Symposium on Networks-on-Chip (2018)
20. Bjerregaard, T., Mahadevan, S.: A survey of research and practices of networkon-chip. ACM Comput. Surv. **38**(1), 1–51 (2006)
21. Furber, S.: On-chip networks. Commun. ACM **49**(9), 43–47 (2006)
22. Chrysos, G., et al.: A survey on network-on-chip architectures and design methods. In: International Conference on Modern Circuits and Systems Technologies (2014)
23. Lee, S., et al.: A tree-based network-on-chip topology for high performance socs. In: Asia and South Pacific Design Automation Conference (2005)
24. Murali, S., De Micheli, G.: Mapping and physical planning of networks-on-chip architectures. In: Design Automation Conference, pp. 27–32 (2004)
25. Guan, N.L., et al.: Exploring wireless network-on-chip architectures using MMWAVE radios. In: International Symposium on Networks-on-Chip (2011)
26. Glass, C.J., Ni, L.M.: The turn model for adaptive routing. In: International Symposium on Computer Architecture, pp. 278–287 (1992)
27. Chien, A.A., Kim, J.: In: IEEE Transactions on Parallel and Distributed Systems, vol. 11, pp. 729–738 (2000)
28. Ababei, C., et al.: Routing with congestion awareness for networks-on-chip. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **28**(2), 237–250 (2009)
29. Wu, J.: LBDR: a fault-tolerant routing scheme for 3d NOCS. In: International Conference on ASIC (2011)
30. Moscibroda, T., Mutlu, O.: A case for bufferless routing in on-chip networks. ACM SIGARCH Comput. Archit. News **37**(3), 196–207 (2009)
31. Goossens, K., et al.: A reliable, real-time NOC for safety-critical systems. Integr. VLSI J. **46**(2), 154–170 (2013)
32. Shacham, A., Bergman, K., Carloni, L.P.: Photonic networks-on-chip for future chip multiprocessors. IEEE Trans. Comput. **57**(9), 1246–1260 (2008)
33. Abdelrazek, A., et al.: Machine learning for network-on-chip design space exploration. In: International Symposium on Networks-on-Chip (2017)
34. Jerger, N.E., Thottethodi, M., Kannan, R.: On-chip wireless communication. Proc. IEEE **106**(1), 54–73 (2017)
35. Shacham, A., Bergman, K., Carloni, L.P.: Photonic networks-on-chip for future generations of chip multiprocessors. IEEE Trans. Comput. **57**(9), 1246–1260 (2008)
36. Benini, L., De Micheli, G.: Multi-technology networks-on-chip. IEEE Design Test **34**(6), 10–17 (2017)
37. Wang, L., Zhang, Y., Qian, D.: Hybrid network-on-chip architectures: design and evaluation. IEEE Trans. Very Large Scale Integr. VLSI Syst. **29**(6), 1105–1118 (2021)
38. Shen, Y., Fu, Y., Zhang, L.: Deep reinforcement learning for adaptive routing in networks-on-chip. IEEE Trans. Parallel Distrib. Syst. **31**(8), 1937–1950 (2020)

39. Abdelhadi, A., Clancy, T.C.: Reinforcement learning for network-on-chip management: a survey. ACM Comput. Surv. **54**(5), 1–36 (2021)
40. Li, Z., Shen, Y., Qian, D.: A survey of machine learning techniques for networkonchip optimization. IEEE Access **10**, 39215–39235 (2022)
41. Kim, D., Lim, S.K.: Through-silicon-via aware interconnect design for 3d integrated circuits. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **18**(7), 1002–1014 (2010)
42. Yan, J., Yang, Y., Wang, Y.: Lifetime reliability-aware routing for networks-onchip. IEEE Trans. Very Large Scale Int. (VLSI) Syst. **22**(8), 1771–1784 (2014)
43. Murali, S., De Micheli, G.: Mapping and physical planning of networks-onchip architectures with quality-of-service guarantees. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **26**(11), 1943–1956 (2007)
44. Coskun, A.K., Rosing, T.S.: Temperature-aware system-level optimization for multicore architectures. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **38**(5), 878–891 (2019)
45. Reagen, B., Shao, U., Wei, G.-Y., Brooks, D.: MachSuite: benchmarks for accelerator design and customized architectures. In: Proceedings of the IEEE International Symposium on Workload Characterization (IISWC), pp. 110–119. IEEE (2016)
46. Parashar, A., et al.: Timeloop: a systematic approach to DNN accelerator evaluation. IEEE Micro **39**(3), 20–29 (2019)
47. Wang, H., Parashar, A., Emer, J.S.: AI-aware network-on-chip design for heterogeneous accelerators. IEEE Micro **43**(2), 34–44 (2023)
48. Fiorin, L., Palermo, G., Lukovic, S., Silvano, C.: Security issues in networks-onchip. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **27**(6), 1116–1129 (2008)
49. Rajendran, J., Karri, R., Sinanoglu, O.: Security analysis of integrated circuit layouts. IEEE Design Test **31**(3), 6–18 (2014)
50. Ancajas, D.M., Chakraborty, K., Tahoori, M.B.: Low-cost and scalable intrusion detection system for networks-on-chip. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **33**(8), 1220–1233 (2014)
51. Gebotys, C.H.: Hardware security and trust. IEEE Design Test **36**(2), 6–13 (2019)
52. Chen, Y.-H., Krishna, T., Emer, J., Sze, V.: Eyeriss: an energy-efficient reconfigurable accelerator for deep convolutional neural networks. IEEE J. Solid-State Circuits **52**(1), 127–138 (2016)
53. Sze, V., Chen, Y.-H., Yang, T.-J., Emer, J.S.: Efficient processing of deep neural networks: a tutorial and survey. Proc. IEEE **105**(12), 2295–2329 (2017)
54. Li, Z., Qian, D., Zhang, Y.: Machine learning-based routing and congestion control for networks-on-chip. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **38**(10), 1896–1909 (2019)
55. Farahnakian, F., Ashraf, A.J., Pahikkala, T., Plosila, J., Liljeberg, P., Porres, I.: Learning-based dynamic power management for multicore systems. ACM Trans. Embed. Comput. Syst. **16**(5), 1–20 (2017)
56. Pande, P.P., Grecu, C., Jones, M., Ivanov, A., Saleh, R.: Performance evaluation and design trade-offs for network-on-chip interconnect architectures. IEEE Trans. Comput. **54**(8), 1025–1040 (2005)