Computer Engineering and Applications Vol. 6, No. 1, February 2017

Network-on-Chip Paradigm for System-on-Chip Communication

Rahmat Budiarto[1], Lelyzar Siregar[2], Deris Stiawan[3]

[1]*Dept. of Computer Information System, Al Baha University, Kingdom of Saudi Arabia*
[2]*Faculty of Comp. Science & Information Technology, Universitas Sumatera Utara, Indonesia*
[3]*Dept. of Comp. Engineering, Faculty of Computer Science, Universitas Sriwijaya, Indonesia*
rahmat @bu.edu.sa, lelyzar.siregar@gmail.com, deris@unsri.ac.id

Developments of modern technologies in electronics, such as communication, Internet, pervasive and ubiquitous computing and ambient intelligence have figured largely our life. In our day micro-electronic products inspire the ways of learning, communication and entertainment. These products such as laptop computer, mobile phones, and personal handheld sets are becoming faster, lighter in weight, smaller in size, larger in capacity, lower in power consumptions, cheaper and functionally enhanced. This trend will persistently continue. Following this trend, we could integrate more and more complex applications and even systems onto a single chip. The System-on-Chip (SoC) technologies, where complex applications are integrated onto single ULSI chips became key driving force for the developments.

One of the main challenges in the SoC development is the power and thermal management [1]. As circuits run with higher and higher frequencies, lowering power consumption is becoming extremely important. Power is a design constraint, which is no more subordinate to performance. Despite process and circuit improvements, power consumption shows rapid growth.

Another challenge is communication architecture [2,3]. Most of current SoCs have a bus-based architecture, such as simple, crossbar-type or hierarchical buses. Unlikeness the scaling of chip capacity, buses do not scale well with the system size in terms of bandwidth, clocking frequency and power due to the following reasons.
As the number of clients grows, the intrinsic resistance and capacitance of the bus also increase. Thus, the bus speed is inherently difficult to scale up.
A bus system has very limited concurrent communication capability since only one device can use a bus segment at a time.
Since every data transfer is broadcast, the bus is inefficient in energy. The entire bus wire has to be switched on and off. Thus, the data must reach each receiver at great energy cost. Despite the fact that improvements such as split-transaction protocols and advanced arbitration schemes for buses have been proposed, these methods cannot address the ultimate problems. To search the future chip capacity, for high-throughput and low-power applications, hundreds of processor-sized resources must be integrated. A bus-based architecture would become a critical performance and power bottleneck due to the scalability problem.

Network-on-Chip (NoC) was proposed in face of those challenges in around year 2001 in the SoC community [4, 5]. On-chip networks are developed on a single chip and designed for closed systems targeting homogeneous or heterogeneous applications.

With the steady advance in electronic downsizing we have reached a point in which we are incapable of increasing processor performance in traditional ways: in one hand, all significant microarchitectural improvements are already in place and, in the other hand, clock frequency has reached the limits imposed by the power and dissipation capabilities of current technology. For this reason, the microprocessor community has been forced to move towards multicore architectures with increasingly high number of processing cores. As the number of cores increases, established on-chip communication infrastructures (buses) start to become a performance bottleneck because a centralised, shared medium results on increases of power consumption and transmission latency as well as in decreases of per-core available bandwidth. For this reason, decentralised infrastructures (networks on chip) are gaining importance and will be essential once a critical number of cores is reached. Many projects attempted to investigate such infrastructures in a holistic manner: covering several areas of opportunity (network topology, router microarchitecture, interfaces) and considering different numbers of merit (power, performance, area, fault tolerance)

As Networks-on-chip (NOCs) are becoming the de facto communication fabric to connect cores and cache banks in chip multiprocessors (CMPs), routing algorithms, as one of the key components that influence NOC latency, are the subject of extensive research. Static routing algorithms consume low cost but unlike adaptive routing algorithms, do not perform well under non-uniform or bursty traffic. Adaptive routing algorithms estimate congestion levels of output ports to avoid routing traffic over congested ports. As global adaptive routing algorithms are not restricted to local information for congestion estimation, they are the prime candidates for balancing traffic in NOCs.

Trough computing beyond a million processors, bio-inspired massively parallel architectures have been considered to provide more adaptability to support heterogeneous applications [6]. There are challenges in this research area, from understanding brain function at the highest level, through automatic mapping of neural models onto a highly distributed computation platform down to architecture optimization and run-time fault and energy management.
To eliminate the delay of routing from the critical path of a router, researchers proposed look-ahead routing, where routing decisions are performed one hop in advance. The delay of the routing being off the critical path, so researchers proposed dynamic routing methods, which are more complex, for better distribution of traffic in the network.

The current and almost certainly the next generation of multiprocessor hardware relies on wired connectivity between many core processors in a hierarchy of cores, chips, boards and racks. At the present time, interconnect between all the components is achieved using high speed serial or bus connections. Performance requirements of NoC infrastructures in future technology nodes cannot be met by relying only on material innovation with traditional scaling. The continuing demand for low-power and high-speed interconnects with technology scaling obliges looking beyond the conventional planar metal/dielectric-based interconnect infrastructures.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Among different possible alternatives, the on-chip wireless communication network is envisioned as a revolutionary methodology, capable of bringing significant performance gains for multicore SoCs. Wireless NoCs can be designed by using miniaturized on-chip antennas as an enabling technology.

We agree there are many configurations where wireless communications will be either faster as well as lower in energy usage than the existing wired connections. Undoubtedly there will also be lots of situations where wired connectivity should be preferred to wireless due to higher concurrent channel utilization, noise resistance and lower latency.

In conclusion, Network-on-Chip paradigm on System-on-Chip emphasises the communication protocols that consider adaptiveness to accommodate changes in the network traffic due to the heterogeneous applications and at the same time optimize network throughput, minimize latency and efficiency in power usage. Bio-inspired routing architectures are the best potential solutions due to their nature in adapting any extreme environment changes.

REFERENCES

[1] V. Raghunathan, M. B. Srivastava, and R. K. Gupta. A survey of techniques for energy efficient on-chip communication. *Proceedings of Design Automation Conference*, June 2003.
[2] T. Claasen. An industry perspective on current and future state-of-the-art in system-on-chip (SoC) technology. *Proceedings of the IEEE*, June 2006, vol. 94, no. 6, pp.1121–1137.
[3] L. Benini and G. D. Micheli. Networks on chips: A new SoC paradigm. IEEE Computer, January 2002, vol. 35 no. 1, pp. 70–78.
[4] W. J. Dally and B. Towles. Route packets, not wires: On-chip interconnection networks. Proceedings of the 38th Design Automation Conference, 2001.
[5] P. Guerrier and A. Greiner. A generic architecture for on-chip packet switched interconnections. In Proceedings of the Design, Automation and Test in Europe Conference, March 2000, pp. 250–256.
[6] M.A.J. Sethi, F.A. Hussin, N.H. Hamid. Bio-inspired NoC fault tolerant techniques using guaranteed throughput and best effort services, The VLSI Integration Journal, 54 (2016) pp. 65–96.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)