The SoC Interconnect Fabric: A Brief History

by Kurt Shuler, On Aug 06, 2013

The high functional integration of system-on-chip designs today is driving the need for new technological approaches in semiconductor design. Anyone who owns a Samsung Galaxy S4, HTC One or comparable smartphone can see the benefits of integrating onto one chip all the computing functions that were traditionally separate, discrete chips on a PC computer motherboard. For next-generation devices, developers are driving even greater computing power, higher resolution graphics, and improved media processing into the integrated SoCs that enable these systems. This high level of integration is causing on-chip communications and transaction handling to become a system constraint within the SoC, limiting the achievable performance of SoCs no matter how optimized the individual CPU, GPU and other IP blocks.

All indications point to ever-higher levels of integration and further SoC advancements in the years to come. This will enable even more functions to be added, making systems more sophisticated, smaller, more power-efficient, and more cost-effective. Yet there is still one critical area of the chip design that needs to be addressed: The on-chip interconnect fabric.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

A history of interconnect technology in SoC designs shows us that the semiconductor industry traditionally lagged in the development of the fabric that ties a chip's individual components together within the chip. Usually an afterthought in the design process, this critical component has fallen behind Moore's Law. Although evolution occurred in fits and starts, the SoC interconnect did not receive the type of research and development focus that CPUs or GPUs received. As a result, chip designs frequently suffer from routing congestion, physical timing closure problems, increased die size, time-to-market delays, missed market opportunities, efficiency losses, bottlenecks, and squandered performance potential.

Only recently has interconnect fabric technology advanced to catch up with state-of-the art CPU and GPU technology. However, by learning from the missteps of the past, the industry can evolve SoC interconnect technology in sync with the rest of the key components, and unlock the technology's full potential. Now, more than ever, as process nodes shrink and as designers cram more IP blocks and transistors onto a die, the on-chip network that enables critical transactions between functional blocks moves to the top of the list of items that unleash the maximum performance and value from SoC design.

Interconnect Fabric History Phase 1: Buses

The history of interconnect technology has three eras. The first era was driven by buses. A processor would perform read and write transactions over the bus to a DRAM memory and, if it used a different address, to other target peripherals. Eventually, other initiators used the bus, too, and arbiters became necessary to alternately grant different initiators access to their requested targets. Many companies owned and developed their own bus interconnect IP. However, 1996 brought about the first de-facto industry standard bus protocol for on-chip interconnects: ARM's Advanced Microcontroller Bus Architecture (AMBA). This ushered in the advancement of IP core interoperability.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

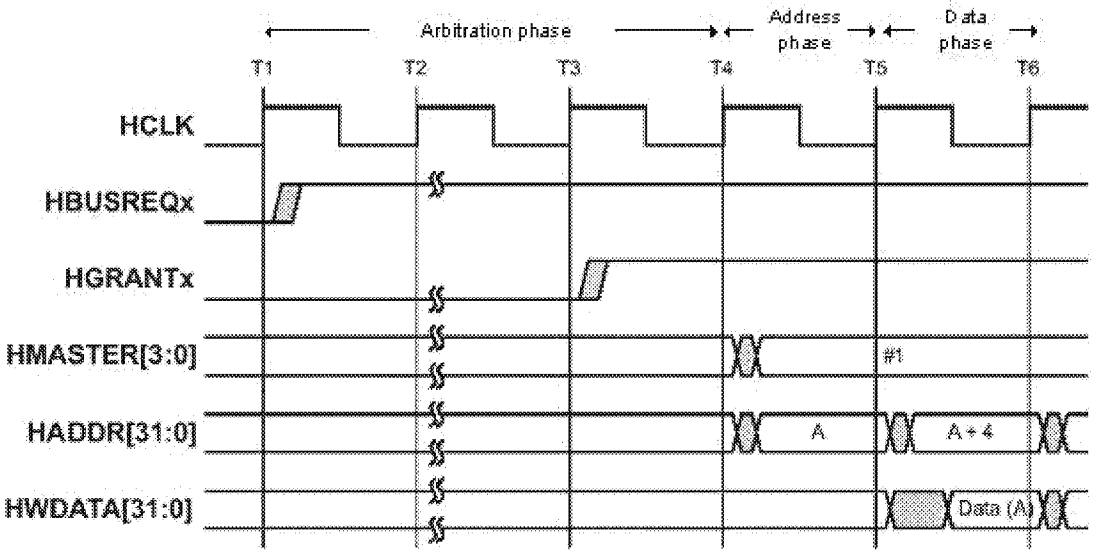


Figure 1: ARM AMBA AHB bus signaling with arbitration, address and data phases of each transfer occurring in different cycles (Source: ARM)

Phase 2: Crossbars

As the integration of multiple cores within chips began in the 1990s, too many initiators trying to access different targets simultaneously created bottlenecks on the bus. Latency issues caused by the long amount of time required for arbitration were a major drawback. The industry wanted to enable concurrent access and more overall system data throughput. A new solution was needed. Crossbars were that solution.

The advantage of crossbars is that connections between initiators and targets communicate simultaneously. Arbitration only happens on a per-target basis, significantly reducing arbitration bottlenecks. However, for initiators that support split transactions, crossbars require complex control logic to keep track of outstanding transactions. Furthermore, they require wide data paths for the initiators and targets that need the most throughput as well as signaling for address and control information. As the number of initiators and targets increased, muxes for wide buses

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

became impractically large. To support continued system scaling, crossbars were cascaded with bridges between fabrics. Bridges between crossbars carried a significant cost in silicon area. The also limited clock frequencies, and added significant latency to transactions.

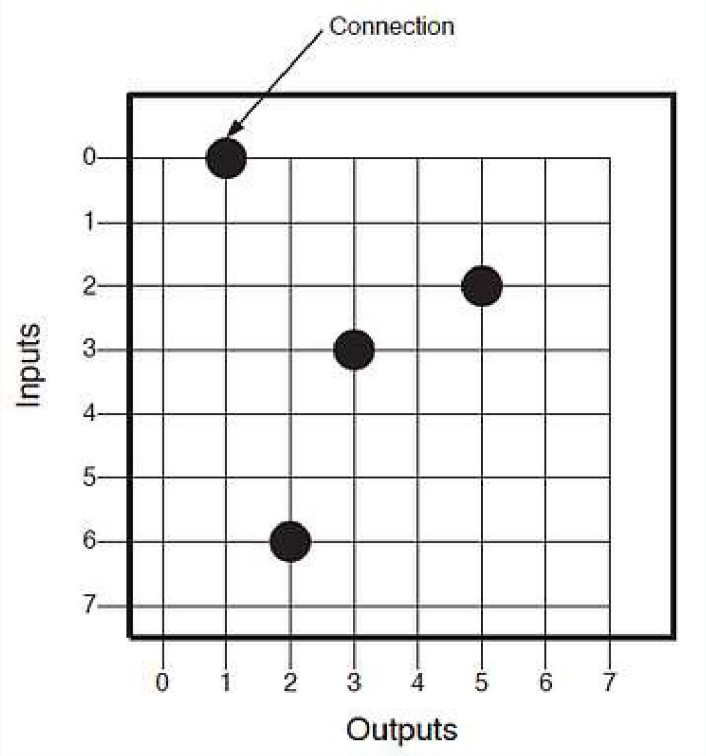


Figure 2: An 8-port unidirectional crossbar switch with four simultaneous connections (Source: Lattice Semiconductor)

As systems-on-chip grew in numbers of IP blocks, busses and crossbars revealed their limitations. Shared busses resulted in contention and hierarchal bus and crossbar designs created complexity. In the third generation, during the 2000s, the initiators and targets became so numerous and widely distributed in the physical floorplan of a chip that the crossbar became a physical wiring nightmare. This stood in the way of efficient place and route and timing closure during the back end of chip design projects.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

Phase 3: Network-on-Chip (NoC)

Packet-based, serialized Network-on-Chip (NoC) technology emerged as a solution to the wiring problem. Packetizing transports address, control, and data information on the same wires, which allows transactions to be sent over smaller numbers of wires while maintaining high quality of service (QoS) for each transmission. Furthermore, distributing the interconnect logic throughout the chip rather than having bridges as chokepoints greatly simplified floor planning of the most complex chips. NoC technology allowed a tradeoff between throughput and physical wires at every point within the on-chip communication network.

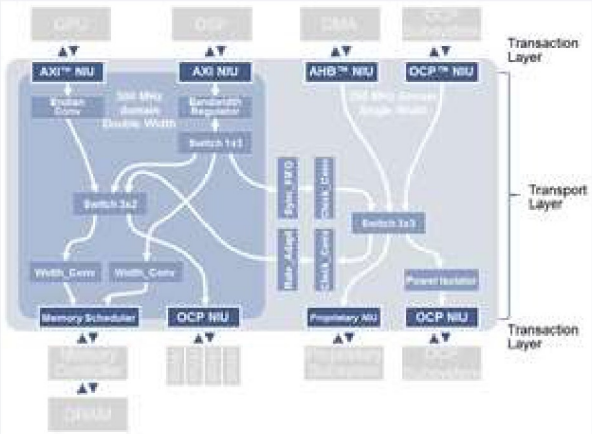


Figure 3: A network-on-chip interconnect with separate transaction, transport and physical layers. (Source: Arteris)

NoC Technology Separates the Transaction, Transport and Physical Layers

Typical interconnect solutions intermingled the transaction, transport and physical characteristics of the interconnect. This historically has made it difficult to deal with changes to the SoC once the interconnect is completed. Comingling these layers also made it difficult to efficiently implement advanced features like QoS, power domains, clock domains, and security.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

NoC Transaction Layer

With Arteris NoC technology, network interface units (NIU) manage communication with a connected IP core. NIUs convert traditional AMBA, OCP and proprietary protocol load and store transactions into packets for transport across the network. At the periphery of the network, NIUs communicate with the attached IP cores via one of a number of standard IP sockets (interfaces), or through custom-built sockets to meet specific customer needs.

Because most of the interconnect logic resides in NIUs that are close to their respective IP blocks, fewer gates are required to place for the interconnect itself.

NoC Transport Layer

NoC technology enables a transport layer to deals exclusively with packets. Only a limited amount of information in packet headers needs to be examined in order to determine the required transport operations. The transport layer can safely ignore the specifics of the transactions being managed at its own level. By doing so, the transport layer has simplified the hardware required for switching and routing functions, and enabled higher operating frequencies.

In short, the interconnect topology is made of simple elements that operate independently without global control requirements. Specific packet handling techniques guarantee quality of service or bandwidth. Optimization can be performed locally on specific routes, without affecting the NoC interconnect as a whole.

NoC Physical Layer

The physical layer defines how packets are actually transmitted between NoC units. Various link types with different capabilities, such as transport links, globally asynchronous/locally synchronous (GALS) links for longer distances, or Chip-to-Chip links, have been employed. Separate transaction and transport layers make it possible to change links, or their characteristics, without affecting the other layers. Because all connections are point-to-point, high fan-out nets are prevented, resulting in improved performance and easier routing.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

Compared to prior interconnect implementations, the Arteris NoC interconnect fabric required fewer connections and wires, was easier for chip designers to use, and simplified timing closure.

NoC technology: The SoC interconnect fabric standard for today, and the future

The largest and most complex chips today utilize NoC on-chip interconnect fabrics and many of these chips have achieved volume leadership in some of the highest growth segments in the industry, such as mobile application processors and LTE modems. This is because NoC technology has resolved the wire routing congesting and timing closure challenges in SoC design, making it easier for designers to create multiple chips based upon a single chip "platform." Its use has allowed developers to eliminate many back-end delays and has accelerated time to market. Using NoC technology also results in smaller chips that have greater throughput, smaller die size, and lower power consumption.

The trend towards greater functional integration within SoCs will drive more and more chip designs to use NoC technology. Today, software programming model simplification is driving a need for cache coherence between initiators within chips. To tackle this challenge, chip designers need to implement cache coherence between initiators that are distributed far and wide around the floorplan of a device. As chips like application processors for mobile devices grow in complexity, area, and transistor count, the need for an advanced interconnect fabric becomes more urgent. Distributed cache coherent interconnect fabrics will be the fourth era in the history of interconnect technology.

Many companies developing SoCs believe they are adding value by developing their own interconnect IP in-house. History shows us that this IP often lags behind the state-of-the-art. History also demonstrates that SoC performance, cost, size and time-to-market suffer when the interconnect technology lags behind that of the CPUs, GPUs, media processing engines and I/O. All of the IP blocks on the chip need to work together for optimum throughput. This decade, however, things are changing as more emphasis is placed on the SoC interconnect fabric.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

—Kurt Shuler is vice president of marketing at Arteris.

Arteris Articles

Discover news and articles about on-chip interconnects, SoC integration automation, and SoC design innovation across the semiconductor industry.

SUBSCRIBE TO ARTERIS NEWS

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

First Name*

mark

Last Name*

zha

Email*

mzhai@nixonpeabody.com

Company Name*

Nixon Peabody

Job Title

IP Attorney

Product Interest *

- [x] FlexNoC Network-on-Chip IP
- [x] FlexWay NoC IP
- [x] Ncore Coherent NoC IP
- [ ] CodaCache Last-Level Cache IP
- [ ] Magillem Connectivity
- [ ] Magillem Registers
- [x] CSRCompiler HSI
- [ ] Harmony Trace
- [ ] Other - IP-XACT Training
- [ ] None specific

- [ ] Get up-to-date information on Arteris products and solutions by opting in to receive emails from us. You can unsubscribe at any time.

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

I have read the Arteris Terms of Use and Privacy Policy. You will receive instant notifications of new blogs via email.

Subscribe

Recent Articles

Arteris Celebrates 3rd Year of Automotive ISO 26262 TCL1 Functional Safety Compliance for Magillem SoC Integration Automation

SemiWiki: NoCs give architects flexibility in system-in RISC-V design

EDN: Extending network-on-chip (NoC) technology to chiplets

Arteris Ncore Cache Coherent Interconnect IP Certified for ISO 26262 Automotive Functional Safety Standard

Arteris Awarded ISO 9001 Certification for Semiconductor System IP Quality Management System

Post By Topic

NoC (1)

SoC (1)

Post By Date

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

November 2023 (7)

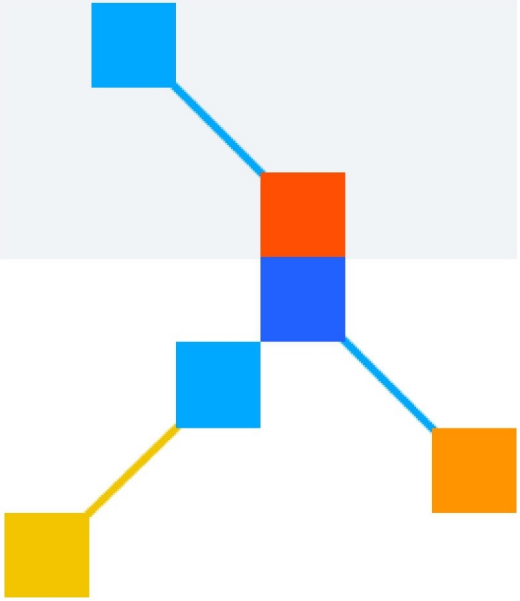
October 2023 (7)

September 2023 (4)

August 2023 (4)

July 2023 (7)

SEE ALL +



ARTERIS IP

Arteris Headquarters

900 E. Hamilton Avenue,
Suite 300
Campbell, CA 95008 USA
+1 408 470 7300

Popular Links

Press Releases

Blog

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338

View office locations →

Events

Customers

Learn

Connect with us

Company

About Arteris

Leadership

Investor Relations

Careers

Contact

Get the latest news, updates, and more from Arteris

I have read the Arteris Terms of Use and Privacy Policy.

Sign up!

Terms of Use **Privacy Policy** **Trademarks & Copyrights** **Trademark Usage Guidelines** **Manage My Privacy** **Do Not Sell/Share My Information**

COPYRIGHT © 2004-2023 ARTERIS, INC

Network System Technologies Ex. 2009
Samsung v. Network System Technologies - IPR 2024-00338