The art of progress is to preserve order amid change and to preserve change amid order.

—Alfred North Whitehead

The goal of this chapter is to provide you information to understand the Intel® QuickPath Interconnect (Intel® QPI) architecture basics to establish a context for studying details in the subsequent chapters. This chapter presents some historical background, including other previous and existing architectures leading to the development of Intel QPI. Specific architectural and electrical features are introduced and connected to the Intel QPI electrical specifications.

Architecture Overview

Intel QuickPath Architecture is designed to provide efficient scalability both logically and electrically. The Intel QuickPath Architecture is a link-based architecture that provides robust point-to-point signaling.
**Historical Context and Motivation**

System architecture based on the Front Side Bus (FSB) is a well established design approach for multiple processor systems. As processor and memory architecture have improved and system speed has increased, the FSB has followed suit to maintain a balanced and performing system. This process has continued to succeed for five generations of Intel processors, starting with the Intel® Pentium® Pro and running through the Intel® Pentium 4 processor families.

This shared-bus architecture, continuing the approach of increasing the bus data rates, has reached intrinsic bus challenges for multiple processor systems. Initially configured with five bus loads, 800 million transfers per second (MT/s) was found to be the limiting data rate. Reducing to three bus loads, and then to two, data rates have been pushed to 1333 MT/s. To continue to achieve greater performance, additional FSB interfaces have been added at a cost of 175 signals each, thus requiring memory controllers to become very large devices with more than 1500 pins.

The FSB is designed to provide communication between multiple processors and a single memory controller. Advantages are deterministic in-sequence events, simpler debugging, and observability. However, the FSB brought with it a performance penalty with each added component to the shared bus. With multiple memory controllers in a system, this bus architecture no longer performs adequately. Memory accesses by multiple components in this configuration are already delayed, and with addition of multiple memory controllers, the FSB becomes over-burdened. To compound this, the FSB shares upstream and downstream signaling on individual signals, but supports only unidirectional signaling, not allowing communication simultaneously in both directions. So the FSB either is sending data upstream or downstream on the shared signal wires. Also, the FSB, with 175 signals, is a heavy burden to component pin counts. Lastly, the Gunning Transistor Logic Plus (GTL+) signaling technology is limited to 1.6 gigatransfers per second (GT/s), which limits the interface bandwidth to 12.8 gigabytes per second (GB/s).

The paradigm needs to change in signaling and logical architecture to get to the next level of driving gigabytes per second between components in the printed circuit board (PCB) and improve memory organization and access latencies. Figures 1.1 through 1.3 show comparisons of example shared bus
systems which use the FSB, a Memory Controller Hub (MCH) and an I/O Controller Hub (ICH). Figures 1.4 through 1.6 show examples of systems using various Intel QPI signaling configurations that use point-to-point links with distributed memory controllers and I/O Hubs (IOH). These sets of figures show how a shared bus and the Intel QuickPath Architecture are in different performance classes.

![Diagram of a computer system with four processors](image)

**Figure 1.1** Four Processor Computer System Based on a Single Front Side Bus

Figure 1.1 shows a system with four processors employing a single shared address and data bus to one chipset component.

- **Key advantages**
  - Simpler bus design, analysis, and debug
  - Global observability among components
  - Deterministic in-sequence events

- **Key disadvantages**
  - Limited bus bandwidth, which goes down with additional components added to the FSB.
Chapter 1: Intel® QuickPath Interconnect Electrical Architecture Overview

Figure 1.2  Four Processor System with Two Front Side System Busses

Figure 1.3  Four Processor System with Four Independent System Busses
Figure 1.2 shows a system with four processors employing two shared address and data busses with point-to-point links to one shared chipset component. Figure 1.3 shows a system with four processors employing four independent address and data busses with point-to-point links to one shared chipset component.

- **Key advantages**
  - Higher bus bandwidth due to point-to-point links

- **Key disadvantages**
  - Loss of global observability
  - Couples system debug with out-of-sequence events
  - Chipsets become the bottleneck for memory bandwidth

---

**Figure 1.4** System with Two Processors Employing Intel® QuickPath Interconnect
Figure 1.5  System with Four Processors Employing Intel® QuickPath Interconnect
Figure 1.4 shows a system with two processors employing Intel QPI links to one chipset component. Figure 1.5 shows a system with four processors employing Intel QPI links to two chipset components. Figure 1.6 shows a system with eight processors employing Intel QPI links to four chipset components.
These figures show some possible configurations of two-, four-, and eight-processor systems respectively that can be built using the Intel QPI (indicated by the red arrow pairs in the figures). Each processor typically has a memory controller on the same die, allowing systems to be more scalable in performance. However this is not essential and systems can have separate, discrete memory controllers. Similarly, the I/O subsystems can either be incorporated onto the same die as the processors or built as separate I/O hubs.

These systems have multiple Intel QPI links connecting the devices to each other. Each one of the links can operate independently of the other links. The performance of such systems can be very high, particularly if the processors are allocated independent tasks, working on data that are optimally distributed across the different memory controllers and close to their own processors. Most current operating systems do recognize such system configurations with Non-Uniform Memory Accesses (NUMA) from each processor and the OS places the data in the memory accordingly.

- **Key advantages**
  - Higher bus bandwidth due to point-to-point links
  - Efficient memory organization and lower memory access latencies
- **Key disadvantages**
  - Loss of global observability
  - Couples system debug with out-of-sequence events

In Figure 1.6 the memory arrays have been left out and the Intel QPI links are depicted as single lines for clarity. The systems shown in Figures 1.4 through 1.6 are fully connected, at least in that each processor has a direct link to every other processor in the system.

The point-to-point links use fewer signals than the shared bus architecture and provides higher bandwidth, transferring data in both directions simultaneously. The Intel QuickPath Architecture also supports multiple memory controllers and efficiently manages the coherence of the processor caches in the system, which is needed for a multiple processor and scalable server systems. The interface provides a robust set of mechanisms to handle errors and recover from them without shutting down the entire system. Intel QPI is defined to meet these needs.
Serial signaling architectures, such as PCI Express† (PCIe), USB, and SATA take advantage of “packet-based” differential signaling to provide higher signal fidelity and protocol flexibility, and some data reliability as well. Intel QuickPath Architecture is constructed to include these capabilities, as well as additional logical and electrical features, to take advantage of both parallel and serial architectures described later in this chapter.

**Intel® QuickPath Architecture Layer Overview**

The Intel QuickPath Architecture is defined to expand the logical and electrical capabilities, as well as to provide flexibility to add to and change the architecture as needed for future requirements. The architecture was also defined to address continuing signaling design and validation challenges. The focus of this book is to address the electrical signaling design, modeling, interconnect implementation, measurement, and validation of Intel QPI. But first, here is an introduction to the overall hierarchy and layer definitions.

The Intel QuickPath Architecture defines different logical layers, with specific functional responsibilities. The Intel QuickPath Architecture layer hierarchy is illustrated in Figure 1.7.

---

Figure 1.7  The Intel® QuickPath Architecture Layer Hierarchy
Intel QPI agents are protocol entities that send/receive a specific type of protocol-level messages. Examples are caching agent, home agent, interrupt agent, etc. Agents are defined at the Protocol layer, and don’t have a Physical layer context. There can be multiple agents within a component. The Protocol layer implements the higher level communication protocol between different Intel QPI agents such as cache coherence (reads, writes, invalidations), ordering, peer to peer I/O, and interrupt delivery. The specific functionality of this layer depends on the platform architecture. The Protocol layer may be bypassed in pure routing agents, which results in low latency transfers from sender to receiver through the interconnection network.

The Routing layer provides a flexible and distributed way to route Intel QPI packets from a specific source to a destination based on the destination agent NodeID. It relies on the virtual channel and message class abstraction provided by the Link layer to specify one or more Intel QPI port and virtual network pairs to route the packet. The mechanism for routing is defined through implementation specific routing tables.

The Link layer separates the Physical layer from the upper layers and provides reliable data transfer and flow control between two directly connected Intel QPI agents. The Link layer provides virtualization of the physical channel into multiple virtual channels and message classes. The virtual channels can be viewed as multiple virtual networks for use by the Routing, Transport, and Protocol layers.

The Physical layer logical sub-block interfaces with the Link layer, and reformats the Link layer data flow control units (flits) to be provided to the electrical sub-block. The logical sub-block also re-formats the electrical sub-block physical data units (phits) into flits to transmit to the Link layer. This logical-layer function is transparent to the electrical sub-block and the Link layer, the Link layer not needing to be aware of the electrical sub-block behavior, and vice versa.

The layered architecture provides contextual separation between layers, which also allows a level of independence to changes within each layer, not affecting the other layer definitions. This contextual independence also applies to the Physical layer logical and Electrical sub-blocks.
Physical Layer Overview

The Physical layer logical and electrical sub-blocks perform separate functions.

The logical sub-block is primarily responsible for Physical layer initialization and training, controlling the electrical sub-block during normal link operation, and Physical layer test and debug hooks. After the Physical layer initialization and training is completed, the logical sub-block works under the direction of the Link layer, which is responsible for flow control. From this link operational point onwards, the logical sub-block communicates with the Link layer at a flit granularity (80 bits) and transfers flits across the link at a phit granularity (20 bits for full-width transfers). A flit is composed of integral number of phits, where a phit is defined as the number of bits transmitted simultaneously in one unit interval (UI) based on the forwarded clock. For instance, a full-width link transmits a complete flit using four phits and receives a complete flit using four phits.

The electrical sub-block defines the electrical signaling technology for high speed data transfer across the link (between agents). Included in the electrical sub-block are the front-end driver and receiver circuits, clock circuitry, analog circuitry for calibrating I/O, detect circuitry, and so on. The electrical sub-block is transparent to the Link layer, and interfaces to the Physical layer logical sub-block.

Architecture Electrical Features and Definitions

Intel QPI electrical signaling establishes significant changes to FSB, which are useful to compare and contrast with PCIe. There are some key similarities, as well as differences in the link and signaling architecture.

Link and Port Definitions

Components with Intel QPI ports communicate using a pair of uni-directional point-to-point links, defined as a link pair. Figure 1.8 illustrates four Intel QPI ports, with the top port expanded to show the individual link elements. Each port comprises a Transmit (Tx) link interface and a Receive (Rx) link interface. For the illustrated example, component A has a Port A that is connected to component B Port B. One uni-directional link transmits from Port A to
Port B, and the other link transmits from Port B to Port A. Please note that this port naming is not dictated by Intel QPI and is only being used to illustrate the Intel QPI link connection between two agents. Also, a “transmit” link and “receive” link is defined with respect to a specific agent or port. For Port A, the Port A transmit link transmits data from Port A to Port B. This same Port A transmit link is the Port B receive link.
Figure 1.8  Four Intel® QPI Ports, Expanding One Port to Illustrate a Link Pair
It is important to note that the link term definition is different than that for PCIe. For PCIe, the pair of Transmit and Receive links is defined as a link, not link pair. One reason that this distinction is made is to provide the Intel QPI link capability to independently change one of the link widths. For example, as shown in Figure 1.9, the Port A Transmit link is a half-width link, while the Port A Receive link is a full-width link. With PCIe, both uni-directional links must remain the same width. There are other logical capabilities that take advantage of this link width independence that are not covered in this book (please refer to the book *Weaving High Performance Multiprocessor Fabric: Architectural insights into the Intel® QuickPath Interconnect* Maddox et al., Intel Press 2009 for more information on this).

![Figure 1.9](image)

**Figure 1.9** Two Different Width Intel® QuickPath Interconnect Links

**Signaling and Clocking Definitions**

Intel QPI ports have two kinds of signals: data and clock. There are no other side-band signals such as interrupts or control signals, simply data and clock signals. The data and clock signals contain the interrupt and control via the packet definitions.