Universal Chiplet Interconnect Express (UCIe)®: An open standard for developing a successful chiplet ecosystem

Dr. Debendra Das Sharma

Intel Senior Fellow, Co-GM Memory and I/O Technologies, Data Center and AI Group, Intel Corporation, Santa Clara, CA 95052, USA
Chair and Co-Founder of UCIe Consortium
debendra.das.sharma@intel.com

I. Introduction

Universal Chiplet Interconnect Express (UCIe)® [1] is an open industry standard interconnect, offering high-bandwidth, low-latency, power-efficient, and cost-effective on-package connectivity between chiplets. It addresses the compute, memory, storage, and connectivity needs across the entire compute continuum, spanning cloud, edge, enterprise, 5G, automotive, high-performance computing, and hand-held segments. UCIe offers a plug-and-play interconnect at the package level, enabling a designer to package chiplets from different sources, including different fabs, using a wide range of packaging technologies.

II. Motivation for On-Package Integration of Chiplets

Gordon Moore predicted the "Day of Reckoning" in his seminal paper where he posited "Moore's law" [1]: *"It may prove to be more economical to build large systems out of smaller functions, which are separately packaged and interconnected."* We are past that inflection point. On-package integration of multiple dies has been widely deployed in mainstream volume CPUs and GP-GPUs [3].

There are many drivers for on-package chiplets. As die sizes keep increasing to meet the growing processing demand, they are exceeding the reticle limit and facing yield challenges in the advanced process nodes. Smaller chiplets connected on package helps designers overcome these challenges.

Lowering the overall portfolio cost with a time to market advantage is a compelling driver for deploying chiplets. For example, the compute cores shown in Figure 1 can be implemented in an advanced process node to deliver leadership power-efficient performance whereas the memory and I/O controller functionality may be reused from a design already deployed in an established process node. Such partitioning also results in smaller dies which results in better yield. It also mitigates IP porting costs which are increasing significantly for the advanced process nodes (Figure 2).

Another value add of chiplets is the ability to offer bespoke solutions. For example, one can choose different numbers of compute, memory, and I/O, and accelerator chiplets depending on the need of the segment. One does not need to do a different die design for different segments, lowering the product cost.

UCIe comprehends these usage models across market segments and is designed to transform the industry.

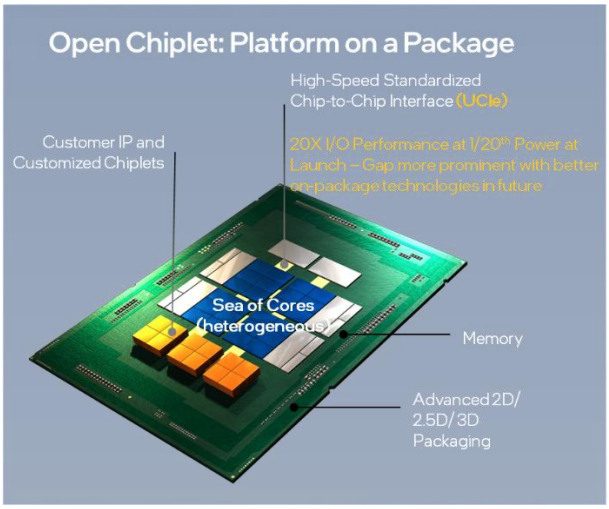


Figure 1: UCIe-based Open Chiplet Ecosystem: Platform on a Package

III. Considerations for an Ubiquitous Industry Standard

Figure 3 represents the necessary ingredients for a standard to be widely deployed, based on our experience in driving successful industry standards such as Peripheral Component Interconnect Express (PCIe)® and Computer Express Link (CXL)®.

An open industry standards body defining a specification is a critical component for wider deployment. Any company can join UCIe and member companies drive its evolution. Thus, companies can invest in the technology without worrying about the viability of the technology, like PCIe and CXL. Compelling key performance indicators (KPIs) catering to a wide range of usages is essential. It is also important to have a complete specification that comprehends all the layers of the stack for ensuring interoperability. As we will see later, UCIe meets these requirements. On-package integration of chiplets has

matured commercially across different manufacturing, assembly, and test companies. This maturity removes another obstacle for adoption.

UCIe is the result of industry leaders with expertise in chiplet integration working together to develop a common standard so that multiple chiplets from different sources can interoperate seamlessly. This will drive a thriving open chiplet ecosystem.

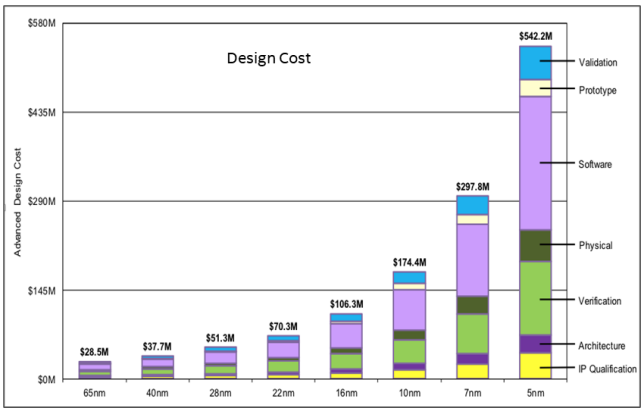


Figure 2: Design cost across different process nodes (Source: IBS, as cited in IEEE Heterogeneous Integration Roadmap)

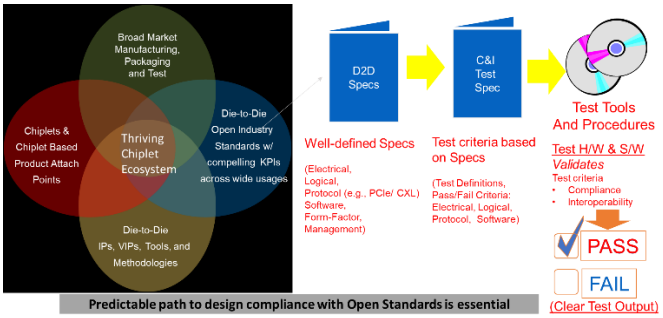


Figure 3: Ingredients of a successful and broad interoperable chiplet ecosystem

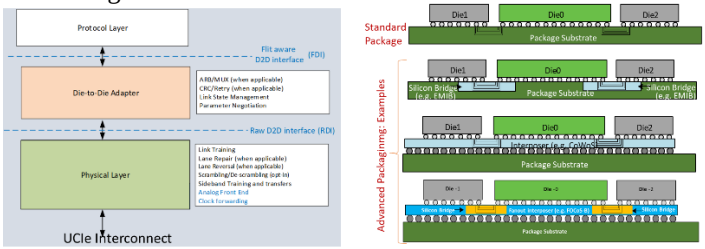
IV. Usage Models and KPIs targeted by UCIe 1.0 Specification

UCIe is a well-specified, layered protocol (Figure 4a). The physical layer is responsible for the electrical signaling, clocking, link training, sideband, etc. The Die-to-Die adapter provides the link state management and parameter negotiation for the chiplets. It optionally guarantees reliable delivery of data through its cyclic redundancy check (CRC) and link level retry mechanism. When multiple protocols are supported, it also defines the underlying arbitration mechanism. A 256-byte FLIT (flow control unit) defines the underlying transfer mechanism when the adapter is responsible for reliable transfer.

UCIe maps PCIe and CXL protocols natively [4] as those are widely deployed at the board level across all segments of compute. This is done to ensure seamless interoperability by leveraging the existing ecosystem where board components can be brought on-package. With PCIe and CXL, SoC construction, link management, and security solutions deployed in today's platform can be seamlessly transported to UCIe.

The usage models addressed by UCIe are comprehensive: data transfer using direct memory access, software discovery, error handling, etc., are addressed through PCIe/ CXL.io; the memory use cases are handled through CXL.Mem; and caching requirements for applications such as accelerators are addressed with CXL.cache. UCIe also defines a "streaming protocol" which can be used to map any other protocol such as a proprietary symmetric cache coherency protocol (e.g., Ultra Path Interconnect). Going forward, the UCIe consortium may innovate on protocols to cover new usage models or enhance existing ones.

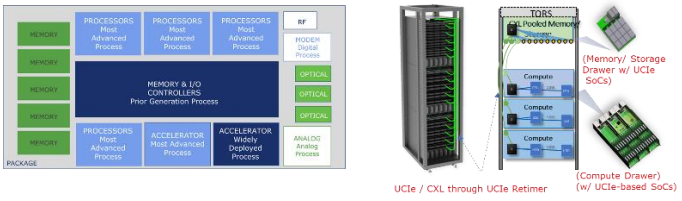
UCIe 1.0 supports two types of packaging, as shown in Figure 4b. The standard package (2D) is used for cost-effective performance. The advanced packaging is used for power-efficient performance. There are multiple commercially available options, some of which are shown in the diagram. UCIe specification embraces all types of packaging choices in these categories.



(a. Layering with UCIe) (b. Packaging Options: 2D and 2.5D)

Figure 4: UCIe : Layering Approach and different packaging choices

UCIe supports two broad usage models. The first is package level integration to deliver power-efficient and cost-effective performance, as shown in Figure 5a. Components attached at the board level such as memory, accelerators, networking devices, modem, etc. can be integrated at the package level with applicability from hand-held to high-end servers with dies from multiple sources connected through different packaging options even on the same package. The second is to provide off-package connecvity using different type of media (e.g., optical, electrical cable) using UCIe Retimers to transport the underlying protocols (e.g., PCIe, CXL) at the rack or even the pod level for enabling resource pooling, resource sharing, and even message passing using load-store semantics beyond the node level to the rack/ pod level to derive better power-efficient and cost-effective performance at the edge and data centers.



(a: Package level Integration) (b: Rack/ Pod level connectivity with UCIe Retimers)

Figure 5: Usage Models supported by UCIe: on-package integration as well as off-package connectivity with different media (e.g., optics, mmWave, electrical cable)

UCIe supports different data rates, widths, bump-pitches, and channel reach to ensure the widest interoperability feasible, as detailed in Table 1. It defines a sideband interface for ease of design and validation. The unit of construction of the interconnect is a cluster which comprises of N single-ended, unidirectional, full-duplex Data Lanes (N = 16 for standard package and 64 for advanced package), one single-ended Lane for Valid, one lane for tracking, a differential forwarded clock per direction, and 2 lanes per direction for sideband (single-ended, one 800 MHz clock and one data). The advanced package supports spare lanes to handle faulty lanes (including clock, valid, sideband, etc.,) whereas the standard package supports width degradation to handle failures. Multiple clusters can be aggregated to deliver more performance per Link, as shown in **Error! Reference source not found.**.

Table 1 summarizes the key metrics for both the packaging options. A die with the standard package design is expected to interoperate with any other design on the standard package. Similarly, a die with the advanced package design will interoperate with any other die designed for the advanced package, even within the wide range of bump pitch from 25u to 55u. It should be noted that the KPI table conservatively estimates performance for the most widely deployed bump pitch today. For example, 45µm bump pitch is used for advanced packaging. The bandwidth density will go up by up to 3.24X if we go with a denser bump pitch of 25µm. Even at 45µm, the bandwidth density of 1300+ GBytes/s /mm or mm^2 (both for linear as well as area) is about 20X of what we can achieve with the most efficient PCIe SERDES. Similarly, PCIe PHY have a power efficiency of ~10pJ/b today which can be lowered by up to 20X with the UCIe based designs due to their shorter channel reach. UCIe also enables for a linear power-bandwidth consumption curve with very fast entry and exit times (sub-ns vs multiple micro-seconds for SERDES based designs) while saving 85+% power. Thus, in addition to being really low power, it also is very effective in power savings, offering compelling power-efficient ultra-high performance. What is important is as the technology advances, these savings would be even more significant. UCIe 1.0 has been defined to meet the projected needs of a wide range of challenging applications through almost the end of this decade.

Table 1: UCIe 1.0 Characteristics and Key Metrics

Characteristics / KPIs	Standard Package	Advanced Package	Comments
Characteristics			
Data Rate (GT/s)	4, 8, 12, 16, 24, 32		Lower speeds must be supported -interop (e.g., 4, 8, 12 for 12G d
Width (each cluster)	16	64	Width degradation in Standard, spare lanes in Advanced
Bump Pitch (um)	100 – 130	25 - 55	Interoperate across bump pitches in each package type across no
Channel Reach (mm)	<= 25	<=2	
Target for Key Metrics			
B/W Shoreline (GB/s/mm)	28 – 224	165 – 1317	Conservatively estimated: AP: 45u for AP; Standard: 110u; Proportionate to data rate (4G – 32G)
B/W Density (GB/s/mm^2)	22-125	188-1350	
Power Efficiency target (pJ/b)	0.5	0.25	
Low-power entry/exit	0.5ns <=16G, 0.5-1ns >=24G		Power savings estimated at >= 85%
Latency (Tx + Rx)	< 2ns		Includes D2D Adapter and PHY (FDI to bump and back)
Reliability (FIT)	0 < FIT (Failure In Time) << 1		FIT: #failures in a billion hours (expecting ~1E-10) w/ CXi Flit Mod

V. Conclusions

The industry needs an open chiplet ecosystem that will unleash innovations across the compute continuum. UCIe 1.0 offers compelling power-efficient and cost-effective performance. The fact that it is an open standard with a plug-and-play model, modeled after several successful standards, and launched by key industry leaders will ensure its wide-spread adoption. We foresee the next generation of innovations will happen at the chiplet level allowing an ensemble of chiplets offering different capabilities for the customer to choose from that best addresses their application requirements.

In the future, we expect the consortium to drive even more power-efficient and cost-effective solutions as bump pitches continue to shrink and 3D integration becomes mainstream. Those may require wider links running slower and get closer to on-die connectivity from a latency, bandwidth, and power-efficiency point of view. Advances in packaging and semiconductor manufacturing technologies will revolutionize the compute landscape in the coming decades. UCIe is well poised to enable innovations in the ecosystem to take full advantage of these technological advances as they unfold.

VI. References

1. D. Das Sharma, “Universal Chiplet Interconnect express (UCIe)®: Building an open chiplet ecosystem”, White paper published by UCIe Consortium, Mar 2, 2022
2. Moore, Gordon, “Cramming more components onto integrated circuits”, Electronics, Volume 38, Number 8, April 19, 1965
3. A. Biswas, "Sapphire Rapids," *2021 IEEE Hot Chips 33 Symposium (HCS)*, 2021, pp. 1-22, doi: 10.1109/HCS52781.2021.9566865
4. “Universal Chiplet Interconnect Express (UCIe) Specification Rev 1.0”, Feb 17, 2022, www.uciexpress.org