UCIe™
Universal Chiplet
Interconnect Express

UCIe 2.0 Specification: Continuing Innovation to Drive an Open Chiplet Ecosystem

Dr. Debendra Das Sharma
Intel Senior Fellow and co-GM Memory and I/O Technologies
Chair of UCIe Consortium

Universal Chiplet Interconnect Express™ (UCIe™) is an open industry standard interconnect offering high-bandwidth, low-latency, power-efficient, and cost-effective on-package connectivity between chiplets. It addresses the projected growing demands of compute, memory, storage, and connectivity across the entire compute continuum - spanning cloud, edge, enterprise, 5G, automotive, high-performance computing, and hand-held segments. UCIe provides the ability to package dies from a variety of sources, including different foundries, designs, and packaging technologies.

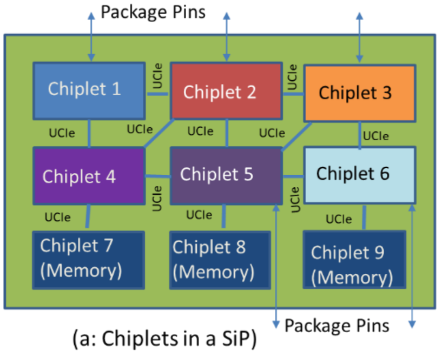
The UCIe 2.0 Specification addresses two broad areas to drive a thriving open chiplet ecosystem. The first addresses manageability, debug, and testing challenges arising in any System-in-Package (SiP) construction with multiple chiplets in a holistic manner. This solution extends beyond the UCIe interface, using UCIe enhancements, in a fully-backwards compatible manner. The second area deals with vertically integrated chiplets with very fine pitches (9 µm down to approximately 1 µm, and potentially lower) using technologies such as hybrid bonding interconnect, which we will refer to as UCIe-3D. We encourage you to read our white papers on UCIe 1.0 and UCIe 1.1 as well as our webinar recordings for additional information on UCIe 1.0 and UCIe 1.1 for insight into previous iterations.

Addressing Manageability, Debug, and Test Challenges at SiP Level Through the Silicon Life Cycle

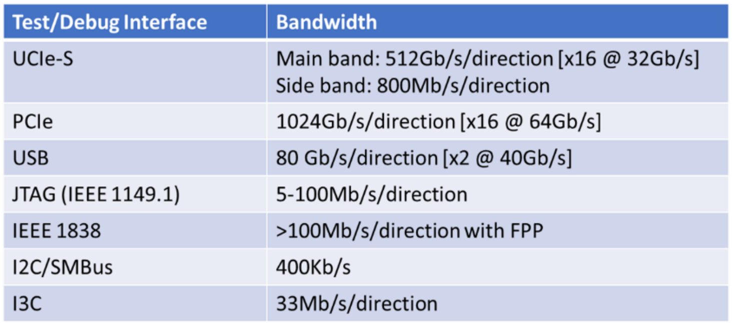
Testability, manageability, and debugging are three main aspects that require continuous innovation. The UCIe 1.0 and 1.1 specifications have several mechanisms in place to deal with various aspects of design for manageability and test/debug/telemetry (collectively referred to as DFx) at the interconnect level. Examples include lane margining, compliance testing, fault reporting, sideband access, and others. However, there are still many challenging problems at the chiplet and SiP level that must be solved to realize the vision of an open, plug-and-play chiplet-based ecosystem. The UCIe Consortium is addressing these challenges holistically, beyond the interface level, to address challenges from the die at sort, through package/bond, to field level – covering the entire silicon life cycle. These enhancements will enable our members to apply these learnings and improve things upstream.

In this paper, we provide examples of the challenges that we need to overcome to realize a broad, plug-and-play, chiplet-based ecosystem.

During die testing at sort, while we can probe bumps, we cannot do the same for micro-bumps; especially as we move towards 25µ bump pitches. So, we must innovate, using other bumps. Similarly, we should be able to manage repair or firmware upgrades seamlessly in the field. Debugging poses a unique challenge for chiplets with limited controllability and observability at the package level (for example, one cannot plug a logic analyzer or scope inside a package). How should the industry handle manageability of chiplets in an SiP? Above all, how do we address these aspects securely? The fact that some chiplets may not be directly accessible from the package pins (see Figure 1a) makes these even more difficult. We are also required to handle a wide range of bandwidth demands. For example, different chiplets have different ranges of bandwidth they need for scan chain, debug, manageability, etc.



(a: Chiplets in a SiP)



Test/Debug Interface	Bandwidth
UCIe-S	Main band: 512Gb/s/direction [x16 @ 32Gb/s] Side band: 800Mb/s/direction
PCIe	1024Gb/s/direction [x16 @ 64Gb/s]
USB	80 Gb/s/direction [x2 @ 40Gb/s]
JTAG (IEEE 1149.1)	5-100Mb/s/direction
IEEE 1838	>100Mb/s/direction with FPP
I2C/SMBus	400Kb/s
I3C	33Mb/s/direction

(b: Bandwidth of various Interfaces: UCIe and External)

Figure 1: (a) Chiplets in an SiP that need to be accessed for various DFx needs, (b) Bandwidth available with various external interfaces as well as UCIe-Standard

Our approach with the UCIe 2.0 Specification is to define a common infrastructure that addresses all of the identified challenges while using existing IP building blocks as well as external interfaces at the package level. We view these features as complementary and our approach works with existing IPs (even non-UCIe IPs) with enhancements to the UCIe PHY. We also employ external package pins to access the chiplets for management, debug, or test through bridging mechanisms that are defined in the specification. These interfaces and IPs must work seamlessly with the on-package UCIe 2.0 links to provide the required external and internal accesses. Figure 1b lists the bandwidth available for different interfaces, offering a diverse menu of choices for SiP designers.

In the UCIe 2.0 Specification, manageability is optional. The mechanisms supported include discovery of chiplets and their configuration; initialization of chiplet structures and parameters (i.e., serial EEPROM replacement); firmware download; power and thermal management; error reporting; telemetry; retrieval of log and crash dump information; test and debug; initiation and reporting of self-test status; and various aspects of chiplet security. These mechanisms leverage existing applicable industry standards and are agnostic to the underlying protocols over chiplets. The mechanisms are intended to work across chiplets from various vendors and support vendor-specific extensions. The capabilities are discoverable and configurable, allowing a common firmware base to be rapidly deployed across SiPs. The required core capabilities of UCIe manageability may be realized through hardware and/or firmware, allowing for enhanced flexibility.

The UCIe 2.0 manageability baseline architecture (Figure 2) defines a bridging function for connecting to an external interface (e.g., SMBus or PCIe®) allowing off-package connectivity. The management fabric in each chiplet consists of multiple management elements, with one of them performing as a management director responsible for discovering, configuring, and coordinating the overall management of the SiP and acting as the manageability root of trust. The UCIe management transport is defined as a media-independent protocol for communication across management entities within a chiplet as well as across chiplets in the SiP. Security mechanisms are defined to provide the desired level of safeguarding depending upon function. Two management link encapsulation mechanisms are defined to transfer UCIe Management Transport packets using side-band and main-band. UCIe defines up to eight independent virtual channels to provide quality of service, each with ordered or unordered semantics. Packets are exchanged based on credits, which are initially negotiated during link training.

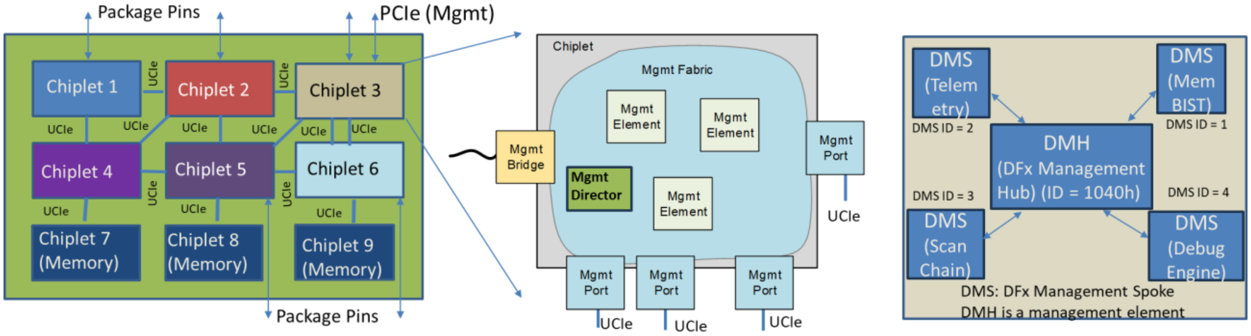
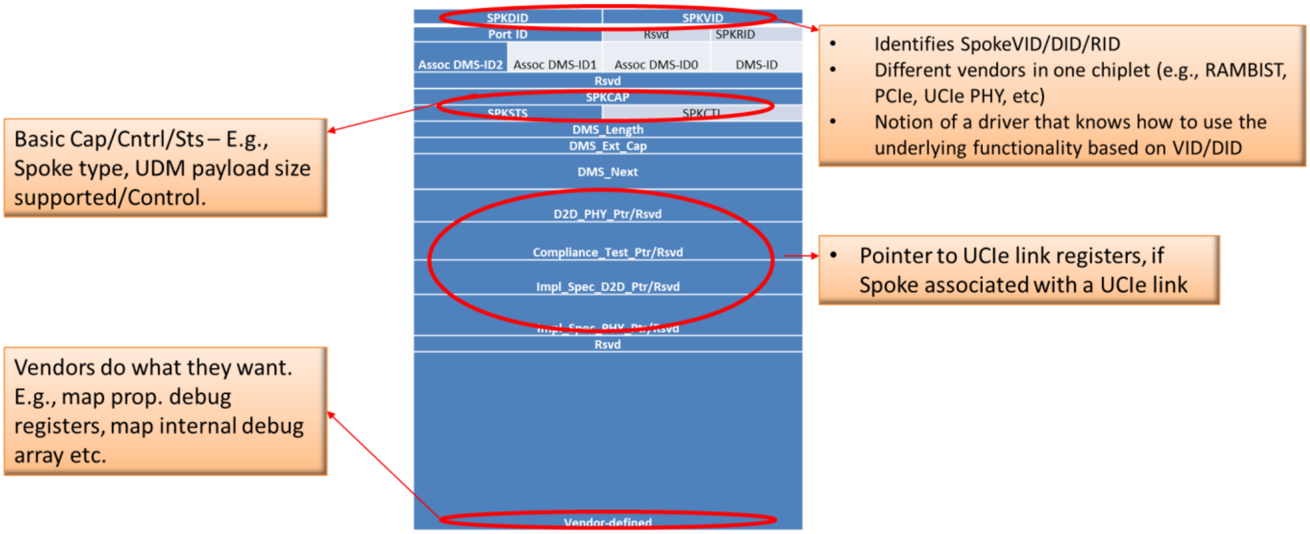


Figure 2: UCIe Management fabric across chiplets, encompassing multiple usage models

The UCIe DFx architecture (UDA) comprehends test, telemetry, and debug and is covered through the management fabric. UDA is based on a Hub-Spoke model within each chiplet (Figure 2). Each chiplet supports a DFx Management Hub (DMH), a management element, that acts as the gateway to access the test, debug, and telemetry capabilities within a chiplet. DMH allows discovery of these capabilities and routs management transport packets relating to these capabilities to various connected DFx Management "Spokes" (DMS). Spokes are the entities implementing a given test, debug, or telemetry functionality. Some examples include, scan controller, MEM BIST, SoC (System-on-Chip) fabric debug, trace protocol engine, core debug, telemetry, etc. Architected configuration registers (Figure 3) with a UCIe-wrapper on top of existing registers provide a common framework for software. For system-level usages, vendor-specific drivers can be loaded to support each unique functionality, depending on (UCIe Consortium-assigned) Vendor ID (VID) and (vendor-assigned) Device ID (DID) of the Spoke. Management Packets for UDA can be sent as either memory access protocol packets (e.g., for discovering the DMH/DMS across chiplets) and/or in vendor-defined UCIe DFx message format (e.g., for sending debug signals over chiplets to package pins such as PCIe to be observed using a Logic Analyzer). Additional usage models are demonstrated in Figure 4.



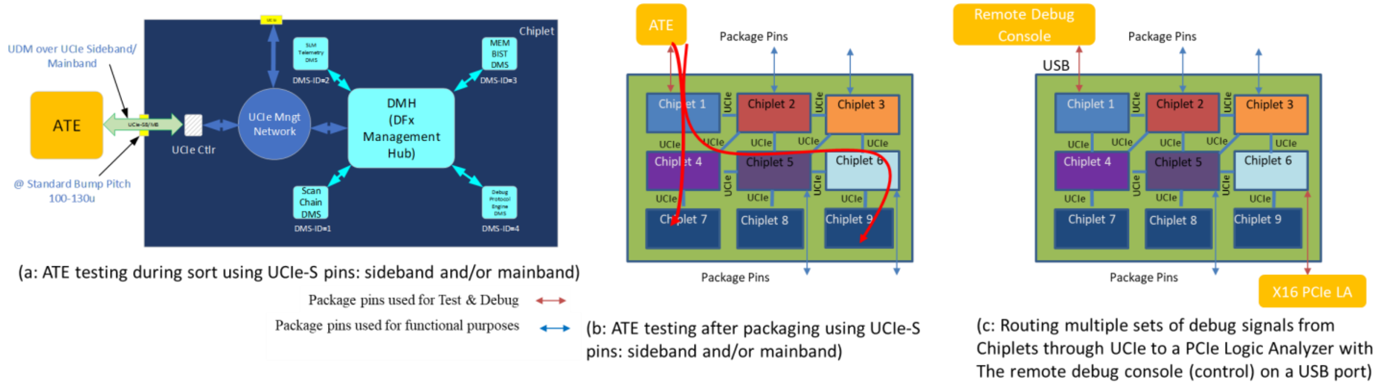
- Identifies SpokeVID/DID/RID
- Different vendors in one chiplet (e.g., RAMBIST, PCIe, UCIe PHY, etc)
- Notion of a driver that knows how to use the underlying functionality based on VID/DID

Basic Cap/Cntrl/Sts – E.g., Spoke type, UDM payload size supported/Control.

- Pointer to UCIe link registers, if Spoke associated with a UCIe link

Vendors do what they want. E.g., map prop. debug registers, map internal debug array etc.

Figure 3: Standardized configuration for UDA for DMS



(a: ATE testing during sort using UCIe-S pins: sideband and/or mainband)

Package pins used for Test & Debug ⟷

Package pins used for functional purposes ⟷

(b: ATE testing after packaging using UCIe-S pins: sideband and/or mainband)

(c: Routing multiple sets of debug signals from Chiplets through UCIe to a PCIe Logic Analyzer with The remote debug console (control) on a USB port)

Figure 4: Example Usages of UDA: Testing at Sort, Package level testing, and Debug

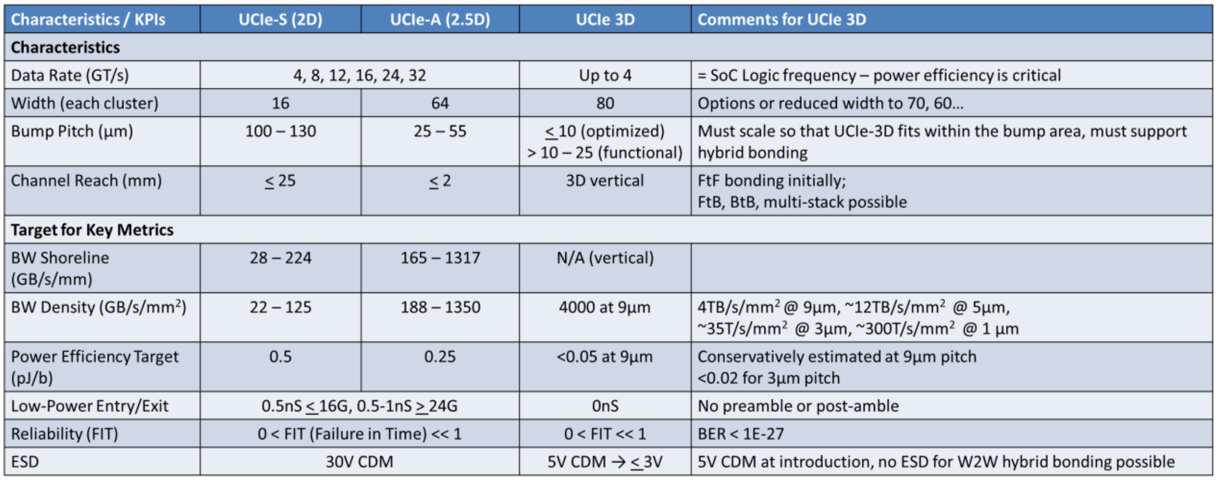
While the management packets can be time-multiplexed on existing UCIe ports, UCIe 2.0 provides the additional capability of adding dedicated UCIe-S ports for manageability and UDA functions. These can be a simple sideband delivering 800 Mb/s/direction with 4 bumps or a half-width (x8), or higher, UCIe-S delivering 256 Gb/s/direction for each x8 at 32 GT/s.

Vertically-Integrated Chiplets for Significant Power Performance Improvements with UCIe-3D

When the UCIe Consortium was formed in March 2022, we released the well-defined UCIe 1.0 Specification addressing planar connectivity (2D and 2.5D). We acknowledged the importance of vertical integration and shared that we intended to work on 3D chiplets. The UCIe 2.0 Specification makes good on that promise with a fully-defined specification encompassing both planar and vertical connectivity.

The technology for delivering vertically connected 3D interconnect chiplets has advanced significantly for more than a decade with commercial offerings of on-package memory and computing, confirming that the demand exists. It is now time to standardize the interface with a menu of options that satisfy the diverse needs found within the ecosystem.

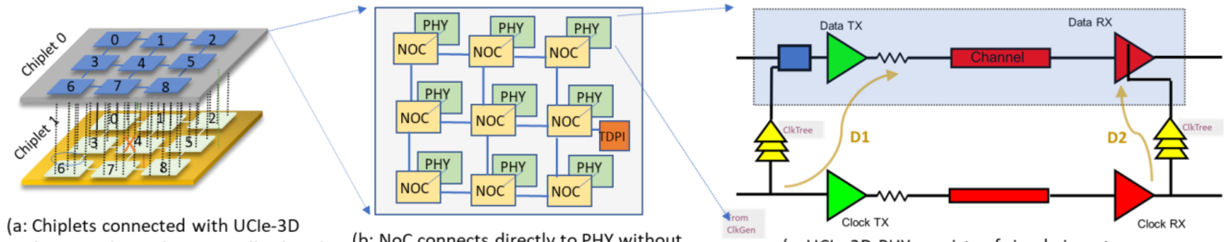
A recent trend for 3D packaging technologies - such as Hybrid Bonding (HB) - has been the aggressive shrinking of the bump pitches between chiplets. UCIe-3D targets bump pitches from 9 µm down to 1 µm, and potentially even lower. 3D interconnect reduces the distance between chiplets to practically 0. Thus, interoperability needs to be constrained to the same bump pitch. While this is not a broad, plug-and-play (i.e., a chiplet with 1 µm bump-pitch can only be hybrid bonded with another chiplet with a 1 µm bump-pitch and not with a chiplet with 9 µm bump-pitch), the key performance indicator (KPI) improvements (such as bandwidth density, power efficiency, etc.) is tremendous. This is illustrated in Table 1.



Characteristics / KPIs	UCIe-S (2D)	UCIe-A (2.5D)	UCIe 3D	Comments for UCIe 3D
Characteristics				
Data Rate (GT/s)	4, 8, 12, 16, 24, 32		Up to 4	= SoC Logic frequency – power efficiency is critical
Width (each cluster)	16	64	80	Options or reduced width to 70, 60…
Bump Pitch (µm)	100 – 130	25 – 55	≤ 10 (optimized) > 10 – 25 (functional)	Must scale so that UCIe-3D fits within the bump area, must support hybrid bonding
Channel Reach (mm)	≤ 25	≤ 2	3D vertical	FtF bonding initially; FtB, BtB, multi-stack possible
Target for Key Metrics				
BW Shoreline (GB/s/mm)	28 – 224	165 – 1317	N/A (vertical)	
BW Density (GB/s/mm^2)	22 – 125	188 – 1350	4000 at 9µm	4TB/s/mm^2 @ 9µm, ~12TB/s/mm^2 @ 5µm, ~35T/s/mm^2 @ 3µm, ~300T/s/mm^2 @ 1 µm
Power Efficiency Target (pJ/b)	0.5	0.25	<0.05 at 9µm	Conservatively estimated at 9µm pitch <0.02 for 3µm pitch
Low-Power Entry/Exit	0.5nS ≤ 16G, 0.5-1nS ≥ 24G		0nS	No preamble or post-amble
Reliability (FIT)	0 < FIT (Failure in Time) << 1		0 < FIT << 1	BER < 1E-27
ESD	30V CDM		5V CDM → ≤ 3V	5V CDM at introduction, no ESD for W2W hybrid bonding possible

Table 1: KPIs for UCIe=3D

The first major benefit of UCIe-3D is increased bandwidth density. This is a two-fold benefit. Firstly, the reduced bump pitch (from 9 µm down to sub-1µm) means the number of wires for a given area increases inversely as a square. For example, comparing 25 µm with 2.5D vs 5 µm in 3D results in a 25X increase in the number of wires in the same area. Second regards the area itself. UCIe-3D offers the advantages of a real connection vs shore-line consumption with UCIe 2D/2.5D. This means that no area is wasted on the PHY at the periphery and the entire chiplet is available for 3D connectivity.



(a: Chiplets connected with UCIe-3D Each Network-on-Chip Controller (NoC) Can connect directly to its neighboring Chiplet NoC using one or more UCIe-3D Linkl(s))

(b: NoC connects directly to PHY without any adapter. A common SoC-level test, debug, Pattern gen/ check, and Infrastructure (TDPI) Block orchestrates training, testing, debug, Etc. across all UCIe-3D Links)

(c: UCIe-3D PHY consists of simple inverters. Data is launched at the rising edge of the clock and Sampled at the falling edge of the forwarded clock. This enables the PHY to scale to sub-micron bump pitches by being bump-limited.)

Figure 5: Chiplets connected vertically with UCIe-3D

Figure 5 shows two chiplets with nine Network-on-Chip Controllers (NOCs) connected using UCIe-3D. To derive the benefits of bump pitch scaling, it is essential to keep the associated circuitry simple, limiting bump. With increased bandwidth density, there is no need to drive higher frequencies. As shown in Table 1, even at 4 GT/s frequency, there are orders of magnitude improvement in bandwidth density over UCIe 2.5D at 32 GT/s (e.g., 300 TB/s/mm^2 with UCIe-3D at 1 μm bump pitch vs 1.35 TB/s/mm^2 with UCIe-2.5D at 25 μm bump pitch). To fit within the reduced bump pitch, we have eliminated the need for (de)serialization, CRC, replay, etc., by choosing the appropriate bit error rate (BER) (as shown in Table 1). Similarly, the ESD protection circuitry must be reduced to 5V CDM initially and eventually eliminated from 3 μm onwards.

The second major benefit with UCIe-3D is lower power usage. With reduced distance (~0), the associated electrical parasitic are reduced. With SoC frequency (<= 4 GT/s), the circuits are uncomplicated - consisting of simple inverters. Combined with reduced frequency, this results in even lower power usage (at least an order of magnitude lower).

Conclusion

UCIe technology is gaining momentum! UCIe Consortium members have announced product developments and provided operational silicon demonstrations based on the UCIe 1.0 and 1.1 Specifications since the Consortium's launch. We are in the early days of a multi-decade journey similar to other successful standards, including PCIe, CXL®, and USB®. Our members are dedicated to making the necessary enhancements to future specifications as we proliferate the technology; UCIe 2.0 is a demonstration of our commitment. The manageability and DFx enhancements signify the on-going commitment to improve existing approaches whereas UCIe-3D denotes our willingness to take on the necessary challenges of delivering exponential improvements in power-efficient performance.

To conclude, I would like to paint a vision of system-in-a-package with several UCIe-3D chiplet stacks connected using existing UCIe-2.5D and UCIe-2D planar interconnects and all the enhancements to come. Today's chiplet integration on-package are like small cities where density is higher than the monolithic chips of a decade ago, which could be likened to small villages. Future SiPs with UCIe-3D will be like a metropolis with skyscrapers, offering very high density. The high density of compute and memory elements closely packaged together means shorter distances for bits to travel resulting in superior performance with decreased power usage. In other words, the future looks very bright, indeed.