



UCle™ 1.1 Specification: Backward Compatible Evolution of UCle for Driving an Open Chiplet Ecosystem with New Usage Models

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

Universal Chiplet Interconnect Express (UCle)[®] 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. UCle provides the ability to package dies from different sources, including different fabs, different designs, and various packaging technologies.

This white paper covers the enhancements included in UCle 1.1 specification. UCle 1.1 specification is fully backward compatible with UCle 1.0 specification. Multiple resources are available regarding UCle 1.0, including a [UCle 1.0 white paper](#), [multiple webinars](#), as well as the [UCle 1.0 Specification](#).

UCle 1.1 addresses four broad areas that encourage a thriving open chiplet ecosystem. These include enhancements for the automotive segment, new streaming protocol usages, cost optimization solutions, and compliance enhancements. Most importantly, these enhancements are fully backward compatible with UCle 1.0 specification, allowing for interoperable evolution of the standard.

Enhancements that Meet Automotive Segment Needs

Automotive is an important segment of the compute landscape. As automotive moves towards using chiplets to leverage the broad ecosystem, it is relying on UCle as the open interconnect to mix and match various chiplets for meeting their needs. To better evaluate automotive requirements and identify appropriate enhancements to the UCle specifications, we have formed an Automotive Workgroup. Based on our investigation, UCle 1.1 specification made the following enhancements:

- **Preventive Monitoring:** The UCle 1.1 specification adds new registers to capture Eye Margin (eye width and height, if applicable) information in a standard reporting format during training. System software can trigger periodic retrain of the link to get eye margin information using the existing UCle 1.0 mechanism.
- **Run-time Testability of Link Health:** We use the periodic parity Flit injection and checking mechanism of UCle 1.0 during mission mode. Any errors detected during the process will be reported in the newly defined UCle 1.1 per-Lane error log register with the ability to send interrupts. System software can then use this information to assess if link repair is needed.

From a field repairability perspective, the existing UCle 1.0 mechanism (e.g., mask faulty lane and retrain) can be used for repair. Thus, no new features are added in UCle 1.1 in this regard.

New and Emerging Usages: Streaming Protocols Can Use the Full UCle Stack

UCle 1.0 supports Streaming Protocol (e.g., AXI, CHI, SMP coherency protocols, SFI, and CPI) only in Raw Mode, as shown in Figure 1. However, given the increasing demand for adopting UCle across the ecosystem and the robustness a die-to-die adapter provides both in terms of enhanced reliability as well as multi-protocol support, we made the following two enhancements with UCle 1.1, while keeping the raw mode support in-tact, as shown in Table 1 below:

- Streaming Protocols can now use the D2D adapter on the FDI interface so that they reuse the CRC, retry, power management features in existing IPs.
- Streaming Protocols are now able to multiplex with other protocols with on-demand interleaving. This enables the co-existence of multiple protocols (e.g., streaming for processing, PCIe for discovery, DMA, TLB, error reporting, interrupt, etc.) for emerging use cases. This leverages the existing UCle 1.0 mechanism of protocol muxing for Streaming Protocol with existing Flit Formats at the FDI interface.

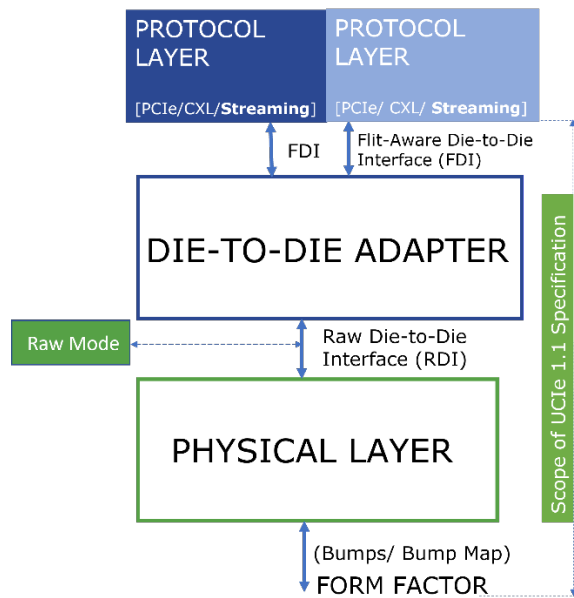


Figure 1: Layering in UCle 1.0 and UCle 1.1

Table 1: Added Support for Streaming with UCle: Format Numbers 2-5 are now supported for Streaming Protocols

Format Number	Flit Format Name	PCIe Non-Flit Mode	PCIe Flit Mode	CXL 68B Flit Mode	CXL 256B Flit Mode	Streaming	
						UCIe 1.0	UCIe 1.1
1	Raw	Optional	Optional	Optional	Optional	Mandatory	
2	68B	Mandatory	N/A	Mandatory	N/A	N/A	Supported
3	Standard 256B End Header	N/A	Mandatory	N/A	N/A	N/A	Supported
4	Standard 256B Start Header	N/A	Optional	N/A	Mandatory	N/A	Supported
5	Latency Optimized 256B without optional bytes	N/A	N/A	N/A	Optional	N/A	Supported
6	Latency Optimized 256B with optional bytes	N/A	Strongly Recommended	N/A	Strongly Recommended	N/A	Supported

Cost Optimization for Advanced Packaging

The UCle 1.0 specification has been defined to interoperate from bump pitches ranging from 55u down to 25u. Interoperability is achieved across these bump pitches by fixing the die-edge length at 388.8 μm, as shown in Figure 2 below. As bump pitches will continue to reduce, the area required goes down as a square. The area efficiency curves in Figure 2 demonstrate that for a fixed shoreline (388.8 μm in the case of UCle-A), one should introduce more columns to take advantage of this square reduction in area (i.e., more bumps along the die edge will reduce the depth). The existing 10-column design in UCle 1.0 is optimal for bump pitches around 45 μm. With UCle 1.1, we added two more arrangements: a 16-column and an 8-column arrangement to effectively stay above the 90% area efficiency shown in the graph in Figure 2. Area efficiency translates to cost and power advantages that will benefit the industry as advanced packaging technology is poised to move towards reduced pitch.

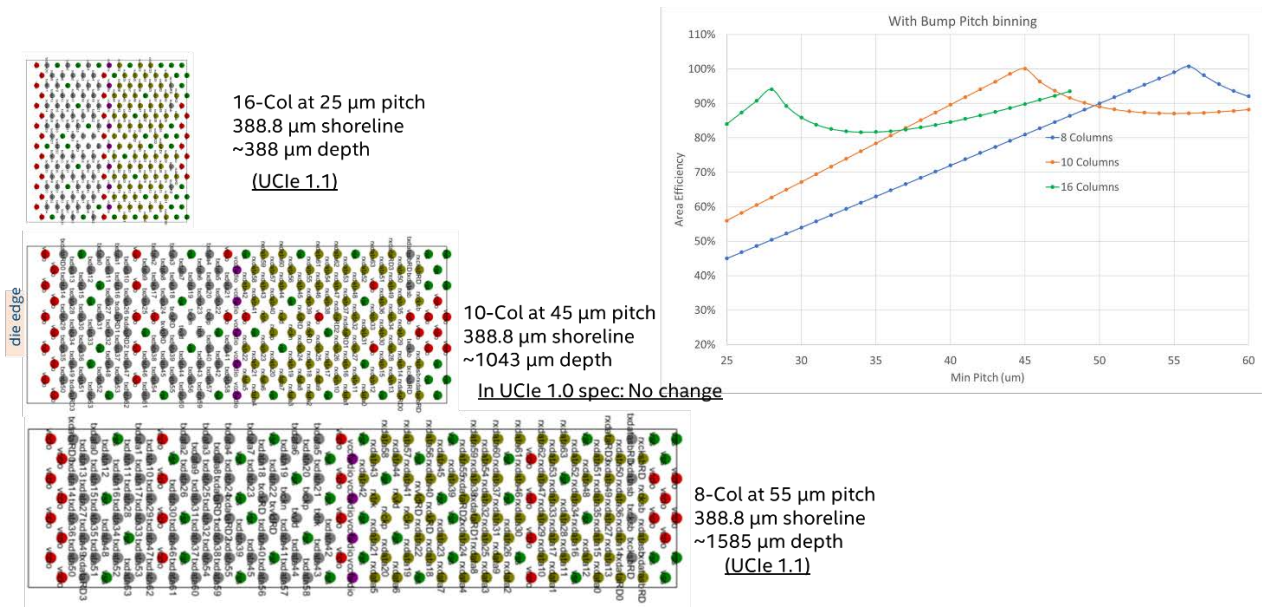


Figure 2: 16-column and 8-column arrangements with UClE 1.1 in addition to the existing 10-column arrangement for UClE-A

In addition to the optimization described above, some evolving usages – such as FPGAs – require narrower width as the building block. In this usage model, there are multiple independent protocol stacks connecting to one or more chiplets based on the bandwidth demand and processing capability. To enable this usage model in a cost-effective manner, we have introduced a x32 native width to realize approximately 40% area savings and enable single layer routing. The x32 can interoperate with a x64 design in degraded mode due to the same 388.8 μm shoreline, as shown in Figure 3.

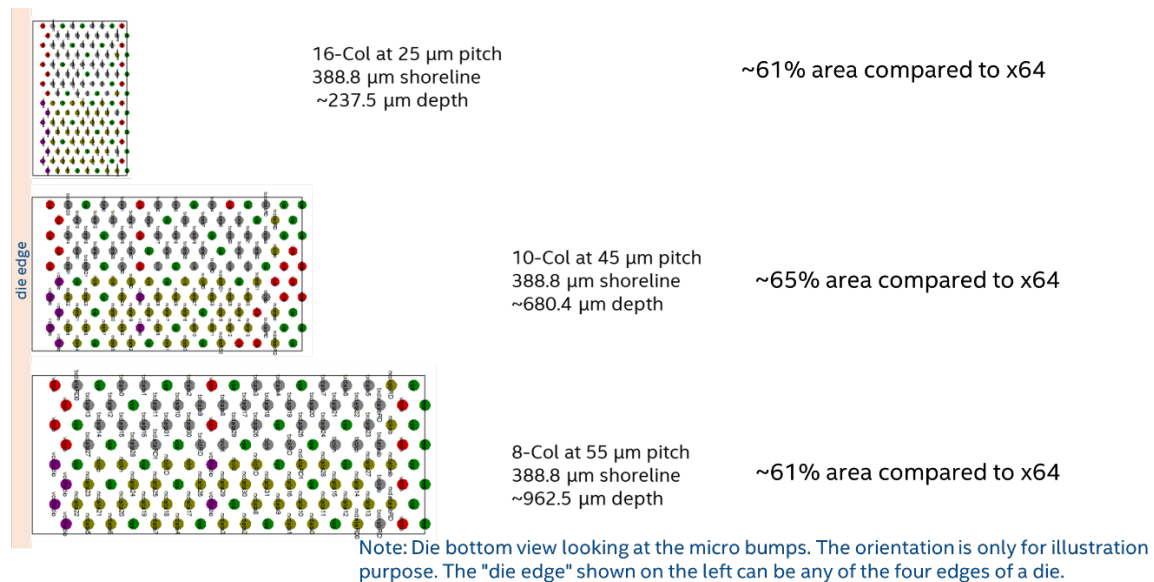


Figure 3: x32 basic building block for UClE-A and the corresponding 16-, 10-, and 8-column layouts

Enhancements for Compliance Testing

Compliance testing is a key attribute for successful standards. UCle 1.0 had hooks for enabling compliance testing which have been strengthened with UCle 1.1. We envision a separate compliance package set-up for advanced vs standard package, as shown in Figure 4 below. We propose testing a golden die by itself first and characterizing the channels. There can be one or more suppliers for this golden die, which will connect through package pins to the external world for controlling the test content. For example, an FPGA may connect to the golden die off-package to run various tests, which in turn can be connected to a laptop for orchestrating the tests and gathering results. The UCle 1.1 specification defines the silicon hooks needed to conduct compliance testing.

There are three areas of compliance, consistent with the three layers. PHY-level compliance deals primarily with signal integrity and link training. To accomplish this, we have defined the mechanisms to conduct timing/voltage margin, bit error rate measurement, lane-to-lane skew, odd/even eye asymmetry, and transmitter equalization, as applicable, with their register-based control. The requirement for the golden die is as outlined above plus the ability to inject errors and cause time-outs in various phases of training to mimic errors that may happen in the field. The die-to-die adapter tests the Flit formation, including NOP and Test Flits, along with CRC and replay, if applicable. These are controlled through configuration registers. The golden die requirements are as outlined above plus the ability to support all the Flit formats, inject errors in Flits, inject errors in the sideband, etc. in a register-controlled manner. The protocol-level compliance is expected to be orchestrated through an FPGA or dedicated silicon connected to the golden die. Details of the protocols can be found within the respective protocols. For example, PCIe® and CXL™ protocols have their respective compliance programs at protocol level that can be used to test the protocol stack of the device under test (DUT).

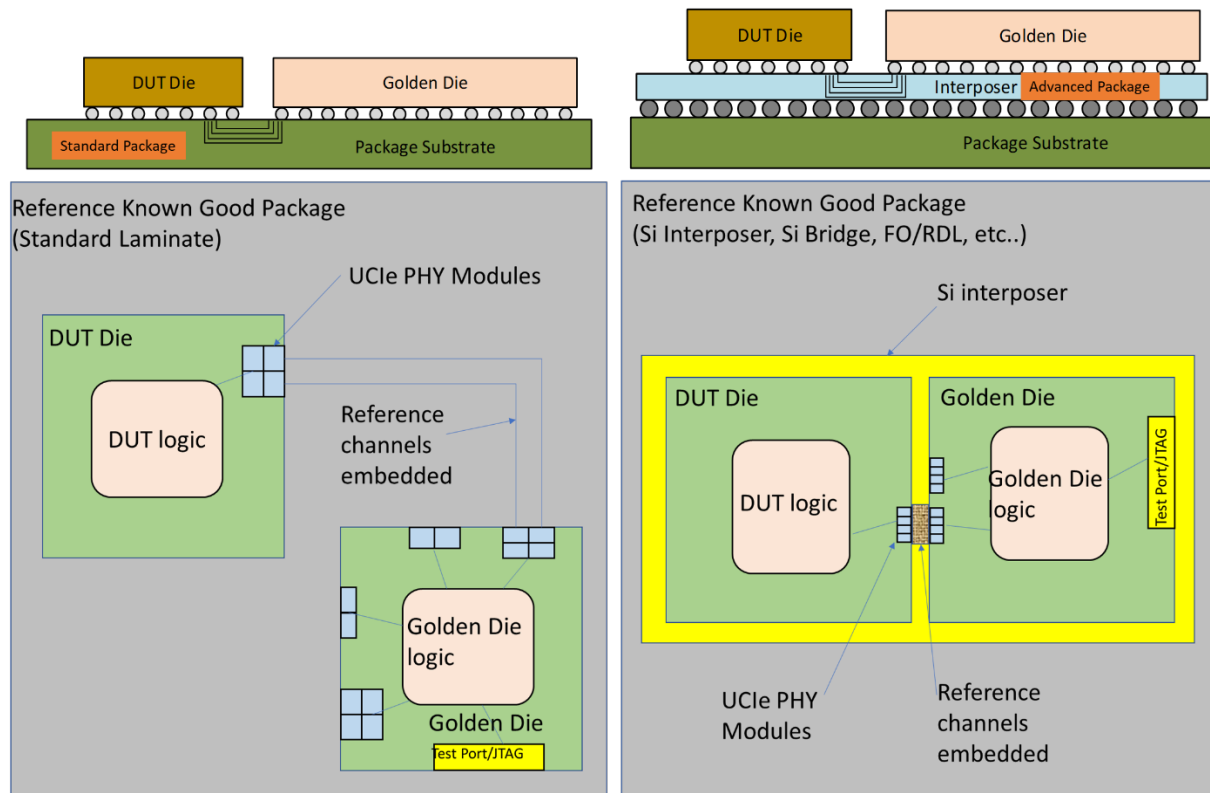


Figure 4: UCle Compliance – both standard and advanced packaging

Conclusions

There is a huge demand for an open chiplet ecosystem that will unleash innovation across the compute continuum. UCle 1.1 builds on UCle 1.0 and offers compelling power-efficient and cost-effective performance. The UCle consortium is committed to moving the technology forward as an open standard with a plug-and-play model, modeled after several successful standards. We foresee the next generation of innovations will happen at the chiplet level allowing customers to choose from an ensemble of chiplets offering different capabilities that best address 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 closer to on-die connectivity from a latency, bandwidth, and power-efficiency perspective. Advances in packaging and semiconductor manufacturing technologies will revolutionize the compute landscape in the coming decades. UCle is well-poised to enable innovation within the ecosystem and take full advantage of these technological advances as they unfold.