CHIPS ALLIANCE

THURSDAY, OCTOBER 14, 2021

Advanced Interface Bus (AIB) for 3D: AIB-3D Version 0.5 Specification Review

Dr. Farhana Sheikh, Senior Engineer, Intel

David C. Kehlet, Research Scientist, Intel

Legal Information

LEGAL DISCLAIMER: Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com. Software and workloads used in performance tests may be optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information, visit www.intel.com/benchmarks. Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Statements in this document that refer to future plans or expectations are forward-looking statements. These statements are based on current expectations and involve many risks and uncertainties that could cause actual results to differ materially from those expressed or implied in such statements. For more information on the factors that could cause actual results to differ materially, see our most recent earnings release and SEC filings at www.intc.com.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

Intel estimated results are based on product specifications. Your costs and results may vary. Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Motivation: 3D ICs and 2.5D D2D PHY Interfaces

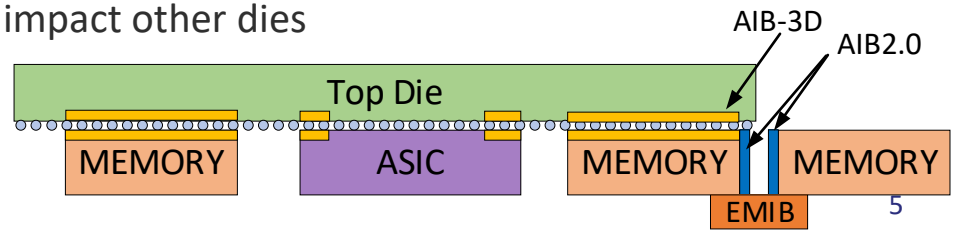
- Increasing use of 3D integration outside of just stacked memories
  - Example: Intel Lakefield (logic-on-logic) [W. Gomes, ISSCC 2020]
- 3D ICs: stacked silicon wafers or dies that are interconnected using through-silicon vias (TSVs) or Cu-Cu connections
- 3D ICs typically designed together by same set of designers or same company
  - Proprietary or simple interfaces, e.g., SRAM memory interface in the vertical direction
- Chiplet ecosystem in vertical direction requires standard interfaces
  - Incorporate into designs to enable 3rd party chiplet integration vertically
- AIB 1.0 and AIB 2.0 are opensource D2D PHY interfaces that enable high-performance interconnect → basis for 3D opensource standard

Outline

- AIB-3D Requirements
- AIB-3D Features and High-Level Architecture
- AIB-3D Channels
- AIB-3D Clocking
- AIB-3D Data Exchange
- AIB-3D Reset and Initialization
- Future Revision Content
- Summary

AIB-3D Requirements

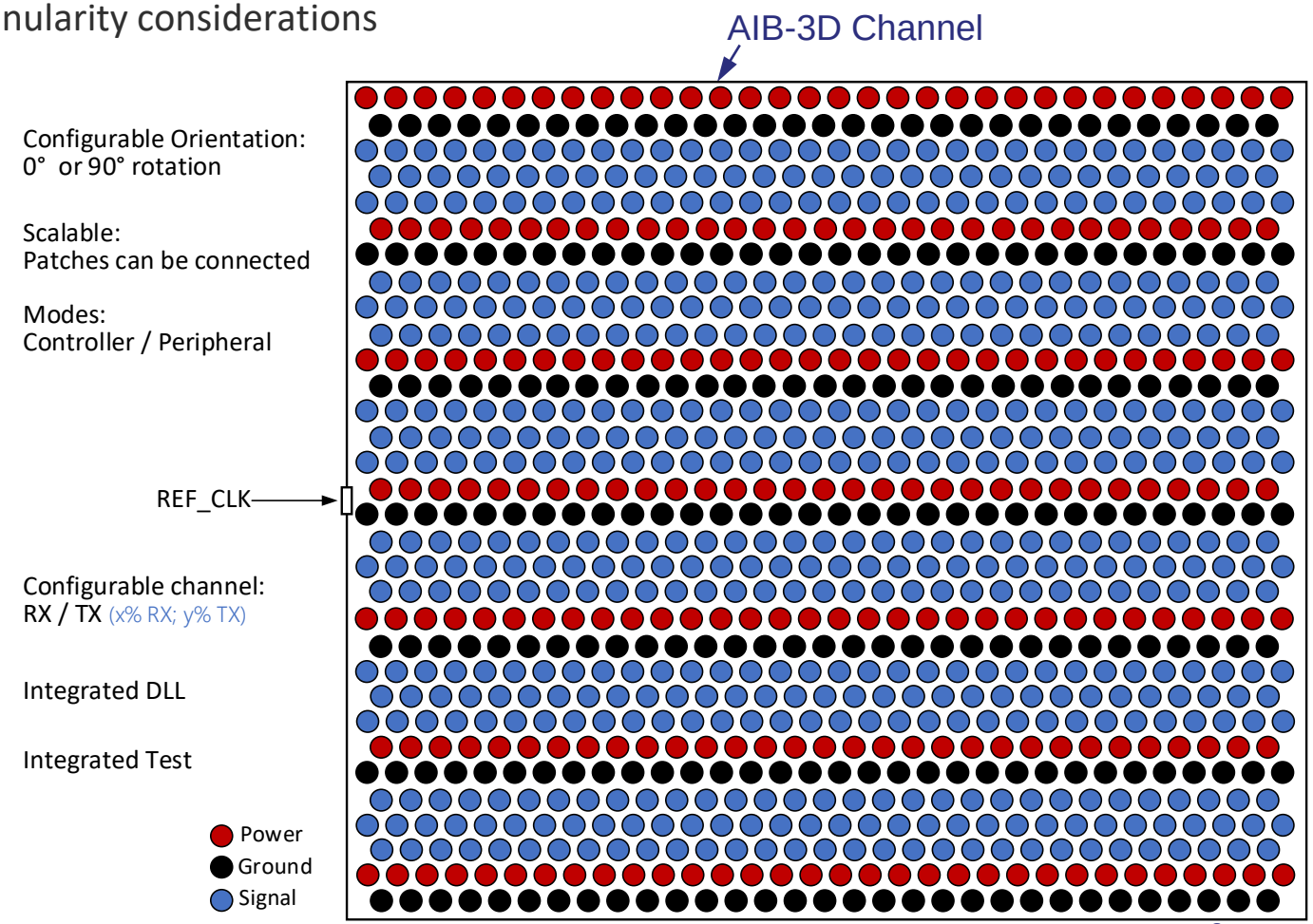
- Scalability across micro-bump and TSV pitch:
  - Future-proof standard across pitch scaling
  - As pitches scale to <10µm, area of patch scales, and require minimal area/power impact of AIB-3D
- SDR vs. DDR support
  - Initially SDR support to keep AIB-3D simple in terms of circuitry and design in over-die integration
  - Future DDR support to allow for high-bandwidth shoreline support
- Half-duplex and full-duplex operation
  - Programmable support for unidirectional and bidirectional data exchange for applications utilizing single port 3D integrated memories or passive dies supporting only unidirectional data exchange
- Point-to-point and point-to-multipoint signaling
  - Initial support for point-to-point signaling and future support for point-to-multipoint: broadcast from one die to multiple die
- Integrated DLL for clock phase alignment to capture data in the center of the eye
- Scalability and reliability: support tiling of AIB-3D channels and redundancy within a channel
- Built-in test and pattern checking for automated verification: 3D integration makes it difficult for test
- Synthesizable or generator based soft-macro for process portability
- Non-intrusive power delivery:
  - Small footprint of AIB-3D channel and intelligent power planning that does not impact other dies
- Compatibility with AIB2.0: support AIB-3D data transfer to AIB2.0



CHIPS ALLIANCE

AIB-3D High-Level Architecture

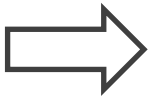
- Wider data bus configured with respect to bump pitch or TSV pitch
  - Examples:
    - 25µm bump pitch ←→ 512-bits TX or RX (half-duplex); 256-bit TRX (full-duplex) [32x32 array, 0.8 x 0.8 mm^2]
    - 10µm bump pitch ←→ 1024-bits TX or RX (half-duplex); 512-bit TRX (full-duplex) [46x46 array, 0.46 x 0.46 mm^2]
    - Data width sized to meet skew requirements, also granularity considerations
- AIB-3D "*patch*":
  - 2Gbps per bump (based on native logic speed)
  - Max. Bandwidth / mm^2: 1600 [25µm bump pitch]
  - Max. Bandwidth / mm^2: 10,000 [10µm bump pitch]
  - Target: 1 to 2 cycle latency per direction
  - Include redundant bumps for repair
  - Integrated test and pattern generation
  - Self-protected I/Os via ESD: 3D drivers are smaller
  - Support for Logic-on-Logic or Logic-on-Memory
  - Initially point-to-point
    - Extension to point-to-multipoint in future
  - 32 x 32 array size (1024 x 25µm bumps)
    (512-bit channel + redundancy + power + control)
  - Array scales with 3D micro-bump / TSV pitch
  - Options:
    - Orientation, modular tiling, modes, configurable channels



25um bump pitch 512-bit AIB-3D Patch 32x32 array
Concept Only – not to scale

AIB-3D Features

Requirement	AIB-3D
Data Rate per bump	**Minimum**: 1Mbps; **Typical**: 2Gbps – 3Gbps
Data transfer	Single Data Rate (SDR)
	Half-duplex and Full-Duplex
IO Voltage Output Swing	0.6V – 0.9V
Electrical and Timing	See Table 10 for Electrical Signal Specifications
Compatibility	AIB2.0



Symbol	Parameter	AIB-3D		
		Near end	Trace	Far end
V_{EH}	Minimum difference between LO and HI (eye diagram height)			TBD
V_{HI}	Minimum output voltage for HI state	0.6 V		
V_{LO}	Maximum output voltage for LO state	0 V		
V_{OS}	Typical output swing	0.9 V		
V_{CM}	Common mode voltage	V_{OS} / 2		
t_{EW}	Minimum duration of valid output (eye diagram width) for data and forwarded clock	TBD UI	TBD UI	TBD UI
t_{BEW}	Minimum duration of valid output (eye diagram width) for receive-domain clock	TBD UI	TBD UI	TBD UI

Features:

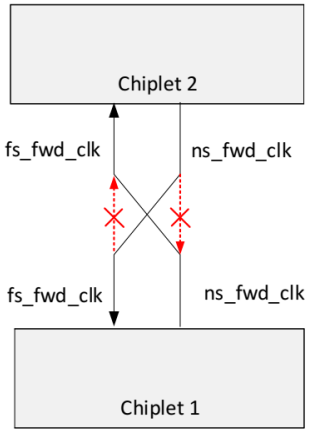
- Typical data rate per bump: 2Gbps – 3Gbps
  - Short distance so kept similar to what may be expected in a monolithic die
  - 512 signals in a single channel → 1Tbps per AIB-3D channel; channels can be tiled for larger bandwidth
  - Serialized high-bandwidth will incur high latency and significant circuit area/power overhead per channel
- Version 0.5 specified for SDR but will enhance in future releases to support DDR
  - AIB-3D DDR envisioned to primarily support high-bandwidth at shoreline
- I/O voltage may differ from core voltage or maybe the same as core: design dependent
- Future possibility: AIB-3D Wireless to support high-frequency chip-to-chip wireless connectivity

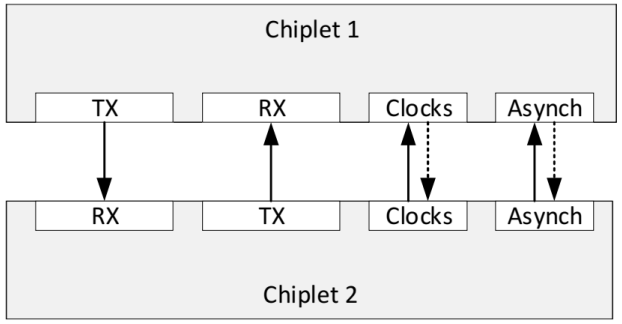
CHIPS ALLIANCE

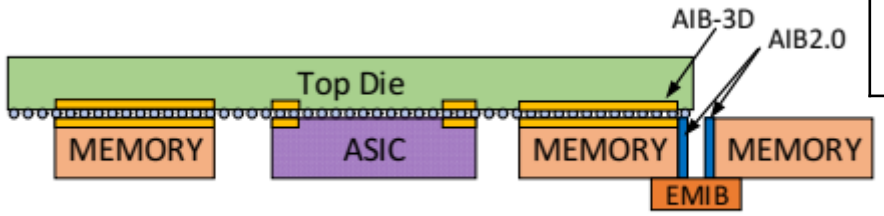
AIB-3D Interfaces

- Use same terminology as AIB2.0
- Can be combined with AIB2.0
- Data, clock, control signals
- New signals include *patch_detect*, *transfer_en*, *transfer_reset*, *patch_reset*

Signal Prefix	Purpose
*ns_*	**Near-side**: chiplet AIB-3D interface being created by designer
*fs_*	**Far-side**: chiplet AIB-3D interface created by a different designer that the Near-side interface will connect to







Signal	Description	AIB-3D SDR	AIB-3D DDR	AIB-3D Wireless
TX	Synchronous data transmitted from the Near-side (Section 2.1.1)	✓	✓	✓
RX	Synchronous data received from the Far-side (Section 2.1.2)	✓	✓	✓
ns_fwd_clk / ns_fwd_clkb	Near-side transfer clock, forwarded from the Near-side to the Far-side for capturing received data (Section 2.1.6.3)	✓	✓	TBD
fs_fwd_clk / fs_fwd_clkb	Far-side transfer clock, forwarded from the Far-side to the Near-side for capturing received data (Section 2.1.6.3)	✓	✓	TBD
fs_transfer_en	Far-side control signal to Near-side to indicate that transfer of data from Near-side can begin; Far-side is Follower	✓	TBD	TBD
ns_transfer_reset	Near-side transfer reset to indicate a new transfer is to start; Near-side is Leader.	✓	TBD	TBD
ns_transfer_en	Near-side control signal to Far-side to indicate that transfer of data from Far-side can begin; Near-side is Follower.	✓	TBD	TBD
fs_transfer_reset	Far-side transfer reset to indicate a new transfer is to start; Far-side is Leader.	✓	TBD	TBD
patch_detect	Used to verify presence of the Leader. Follower output to Leader.	✓	TBD	TBD
patch_reset	Control signal from Leader that is input to the Follower that indicates whether a chiplet has completed power-on reset.	✓	TBD	TBD

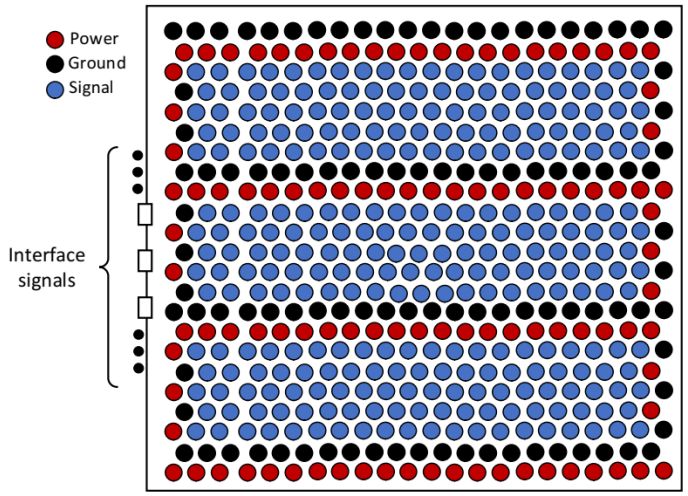
CHIPS ALLIANCE

AIB-3D Channels

- AIB-3D single channel = AIB-3D patch
  - 2-dimensional array (e.g., 512-signal I/Os for 25µm bump pitch in 0.64mm^2)
  - Size is kept <1mm^2 to reduce clock skew across patch and limit area impact to die
- Channels can be tiled

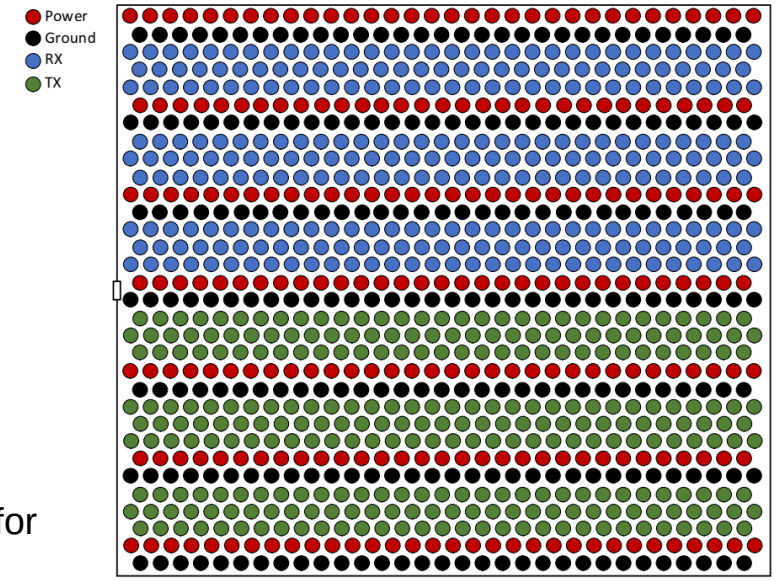
Micro-bump pitch (µm)	Range of data signals per channel (Increments)					
	TX			RX		
	Balanced	All-TX	All-RX	Balanced	All-TX	All-RX
25	16 – 256 (16)	16 – 512 (16)	0	16 – 256 (16)	0	16 – 512 (16)
10	16 – 512 (16)	16 – 1024 (16)	0	16 – 512 (16)	0	16 – 1024 (16)

AIB-3D Channel or "*AIB-3D Patch*"



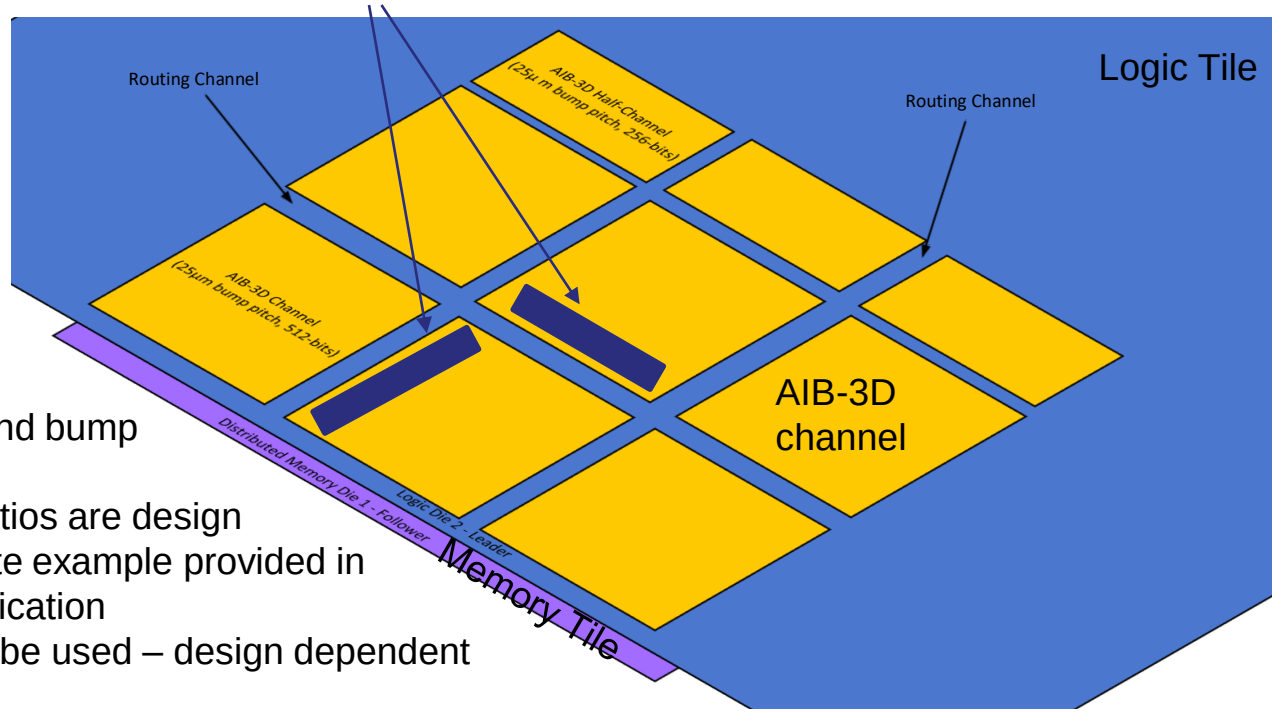
25um bump pitch 512-bit AIB-3D Patch
Concept Only – not to scale
Interface signals may be placed in alternate locations
Power/ground signals may be placed in alternate locations

Balanced full-duplex



25um bump pitch 512-bit AIB-3D Patch 32x32 array
Concept Only – not to scale

- Bumps placed over entire area for mechanical integrity (uniformity)
- I/O placement is design dependent

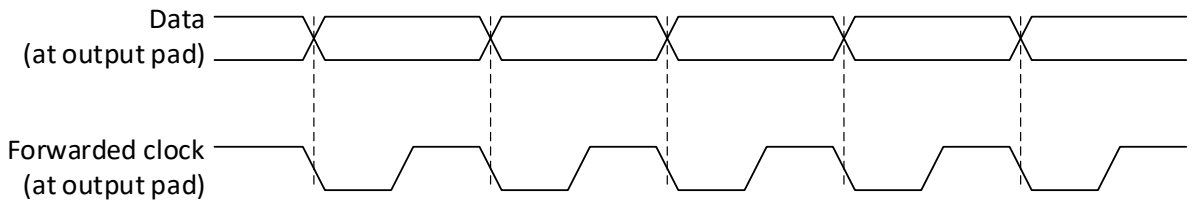


- Tiling requires design dependent power/ground bump planning
- Power/ground bump ratios are design dependent and concrete example provided in future release of specification
- Routing channels may be used – design dependent

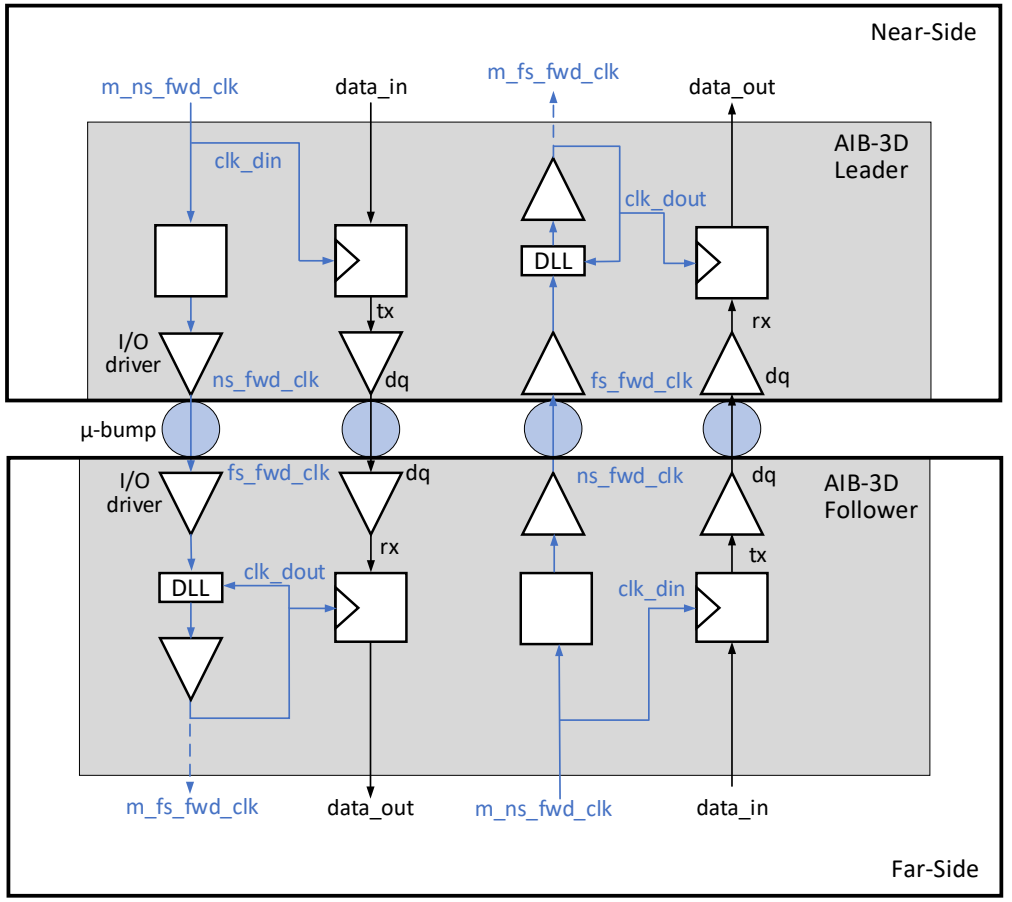
AIB-3D Clocking

- Near-side connects to Far-side using forwarded clock and data as in AIB2.0
- Choose to use forwarded clock from Near-side on return path or use Far-side forwarded clock on return path as below
  - Dependent on design: base and top die have large PVT differences then use two clocks
- DLL recommended
  - Phase correction
- "*ns_fwd_clk*" may be used as sole clock source
- Half-duplex clocking: TBD
- Lowest round trip latency achieved by using *ns_fwd_clk* on return path (see backup)
- Differential clocks maybe used by designers but once converted to single-ended, clock can be used by chiplet and AIB-3D
  - Supported to be compatible with AIB2.0 specification

Latency path	AIB-3D
Transmitting	1.5 – 2 CLK cycle (1.5 – 2 UI)
Receiving	1.5 – 2 CLK cycle (1.5 – 2 UI)

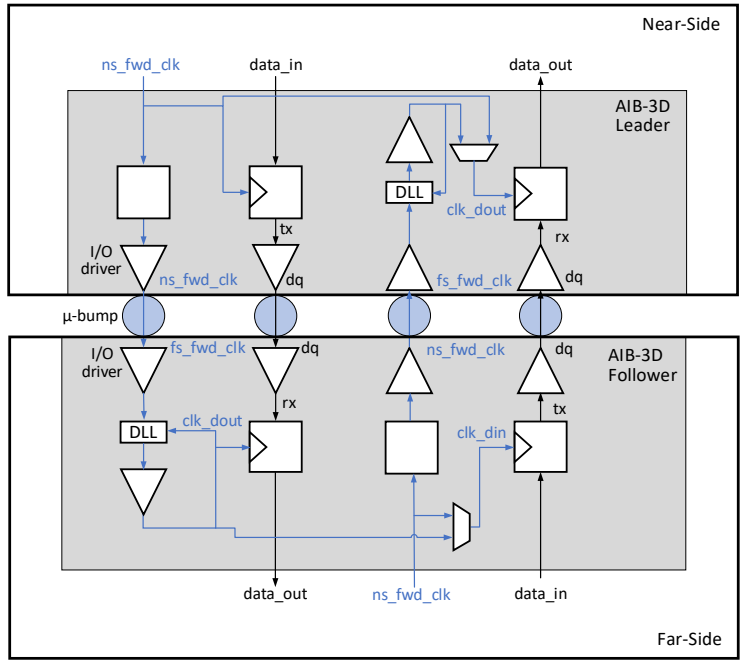


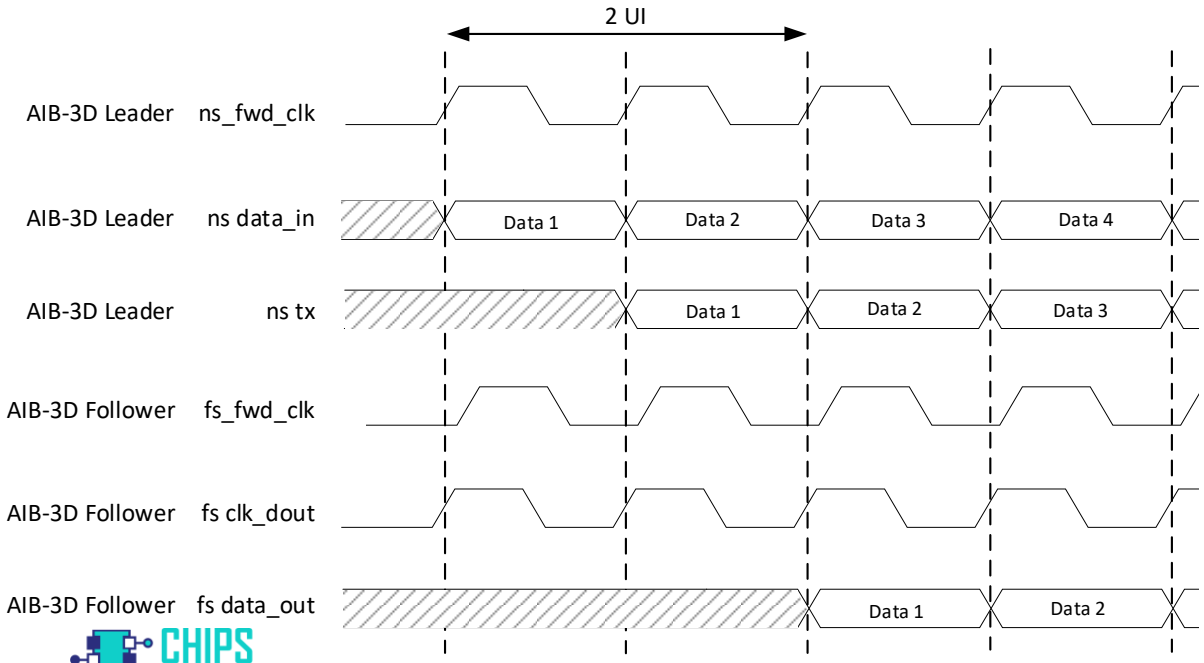
SDR supported in initial release



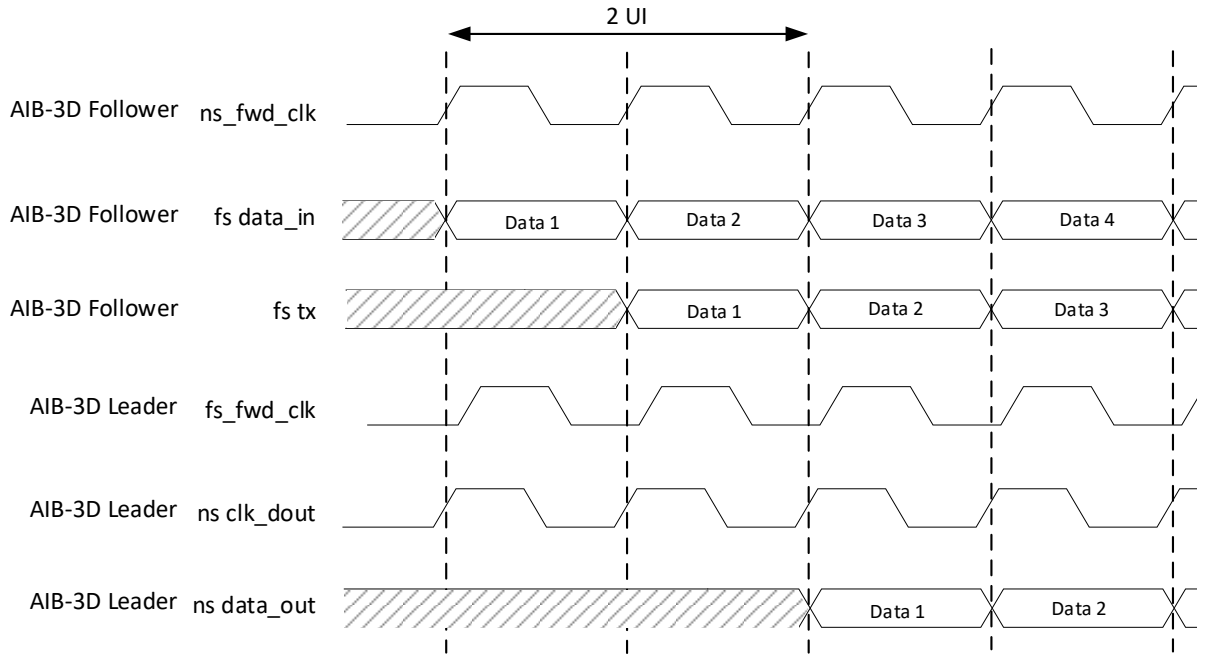
AIB-3D Leader-Follower-Leader Round Trip

- AIB-3D SDR in initial release; future support for AIB-3D DDR
- 2-cycle latency: registered inputs and outputs
- Similar skew requirements for clocking as AIB2.0
- *transfer_en* behaves as a data-transfer ready signal:
  - Deasserted to allow intentional change in clock frequency
  - Receipt of bad data
  - Loss of DLL lock
  - Initiation of reset





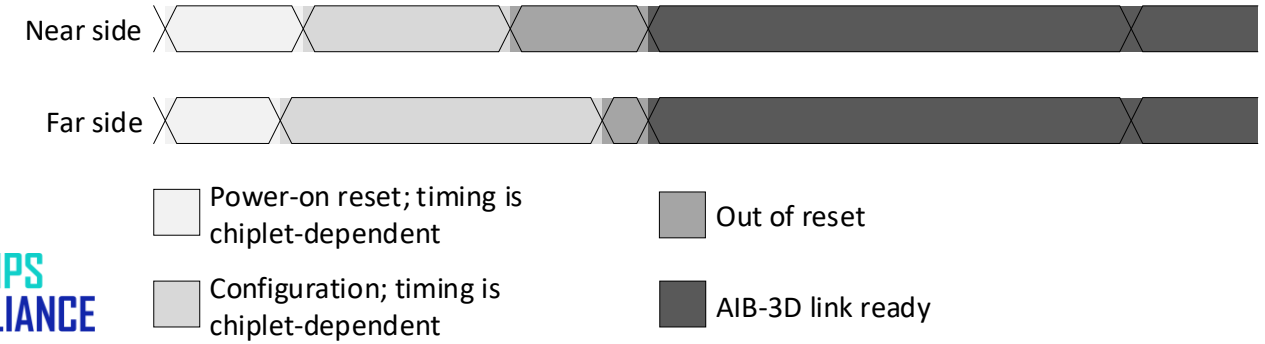
(a) TX path

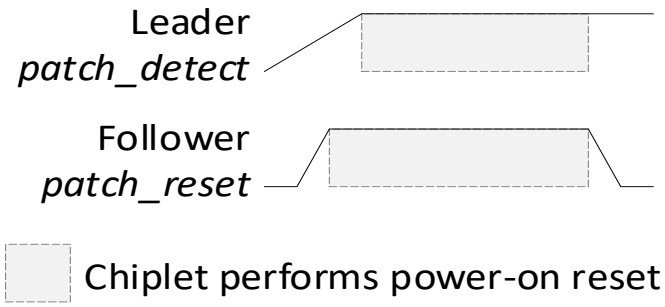


(b) RX path

AIB-3D Reset and Initialization

- Initialization consists of (a) power-on reset synchronization, and (b) calibration
- Multiple AIB-3D interfaces on a single chiplet come out of configuration at the same time but complete reset and calibration at different times depending on implementation
- Data transfer is reset on Near-side and Far-side using the *transfer_reset* signals to reset the AIB-3D channel.
  - All transfers are halted, and input/output registers (flops) are reset
  - Data outputs are placed in standby mode and clocks outputs are put into standby mode
- *patch_reset* and *patch_detect* are used during power-on: behavior dependent on leader/follower
- *patch_reset* when de-asserted by follower indicates configuration can begin
- *patch_detect* when asserted by leader indicates presence to follower





Future Revision Content

- Timing verification based on proposed implementation example
- Built-in test: leverage IEEE 1838 standard as much as is possible
- Physical signal arrangement
- Bump map planning and recommendations
- Configuration details
- Calibration details
- AIB-3D DDR

Summary

- Increasing use of 3D integration outside of just stacked memories
- Chiplet ecosystem in vertical direction requires standard interfaces
- Initial version of draft specification released
  - Require your feedback for improvements
  - Any specific use cases that may not be easily covered by specification?
- Covered important aspects in first version
- More additions to come in future

CHIPS ALLIANCE

BACKUP

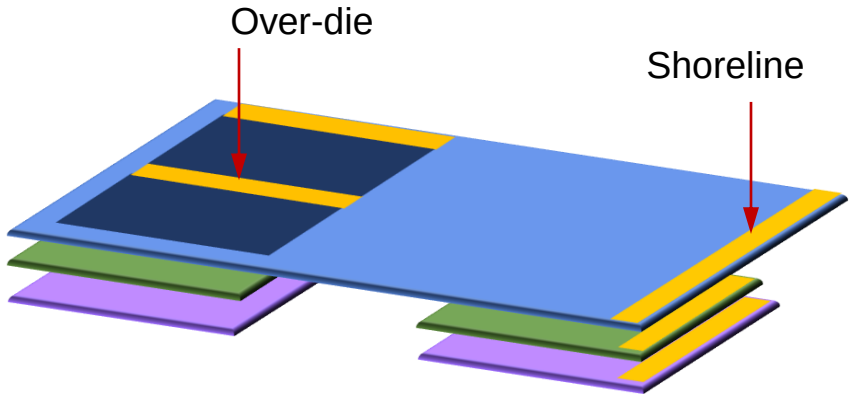
3D PHY Standards: Existing and Requirements

Existing 3D PHY interfaces:

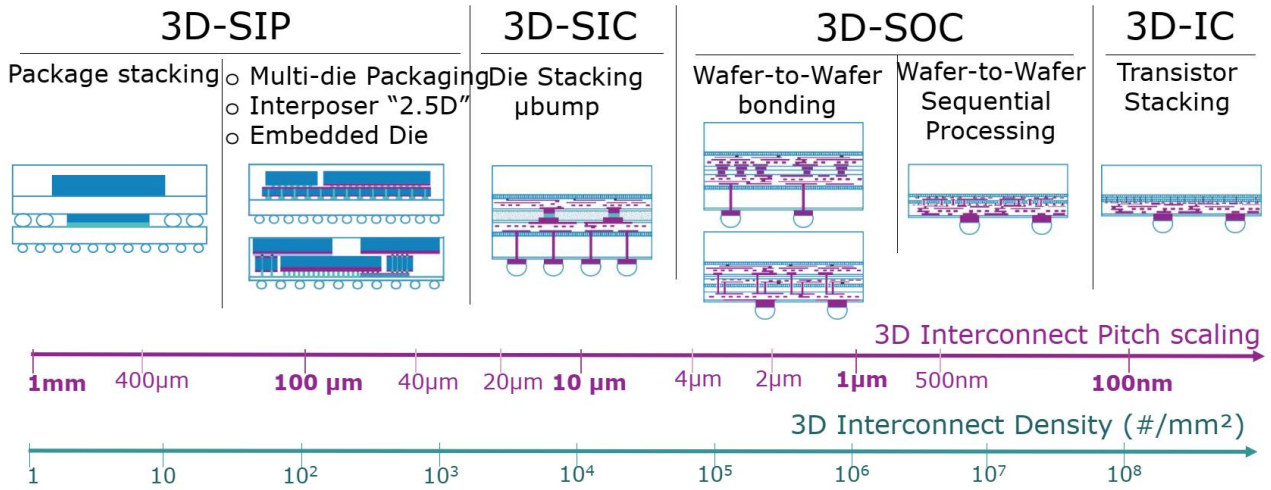
- HBM2, HBM2E, HBM3, Glink-3D [1]
  - Memory-centric or early stages and not open-sourced

Requirements:

- Scalable across wide range of micro-bump and TSV pitch: future proof
- Single data rate
- Half-duplex or full-duplex operation - programmable
- Point-to-point or point-to-multipoint
- Can be used for Logic-on-Logic or Logic-on-Memory
- Integrated PLL/DLL
- Scalable number of channels
- Lane repair
- Built-in test and pattern checker
- Synthesizable or generator based soft-macro
- Non-intrusive power delivery



Reference: AIB-3D Concept, F. Sheikh, Intel Corporation



Reference: E. Beyne, imec, Forum 5.8, ISSCC 2021

CHIPS ALLIANCE

Ref. [1]: J. Bausch, May 24, 2021, EE Times

Some 2.5D D2D PHY Standards

Best possible energy-efficiency

Serial Link	Signaling	Technology	Channel	Reach (mm)	Data Rate (Gbps) Per Wire	Energy (fJ/b)	Note
Ground referenced signaling	SE	28nm	On-chip	1.5	16	255	1-hop used as proxy for interposer
Voltage-Mode Tx with Passive EQ	SE	28nm	Interposer	2.5	20	300	Does not include clocking power
				3.5	18	320	
TSMC LIPINCON	SE	16nm	InFO	0.55	2.8	424	0.3V I/O voltage
	SE	7nm	Interposer	0.5	8	560	
Intel AIB 2.0	SE	-	Interposer, MCM	10	4.0 - 6.4	500	0.4V I/O voltage, 45um / 55um EMIB support
Bunch of Wires (BoW)	SE	14nm	MCM / Interposer	10	5	400	0.7V I/O voltage
			MCM	50	32	700	Bi-directional
MediaTek MLINK	SE	7nm	InFO, CoWoS	<1	12.8 (InFO), 20 (CoWoS)	460 - 500	48um pitch InFO, 40um pitch CoWoS

References: Forum 5, ISSCC 2021; VLSI 2021

AIB-3D based on AIB2.0

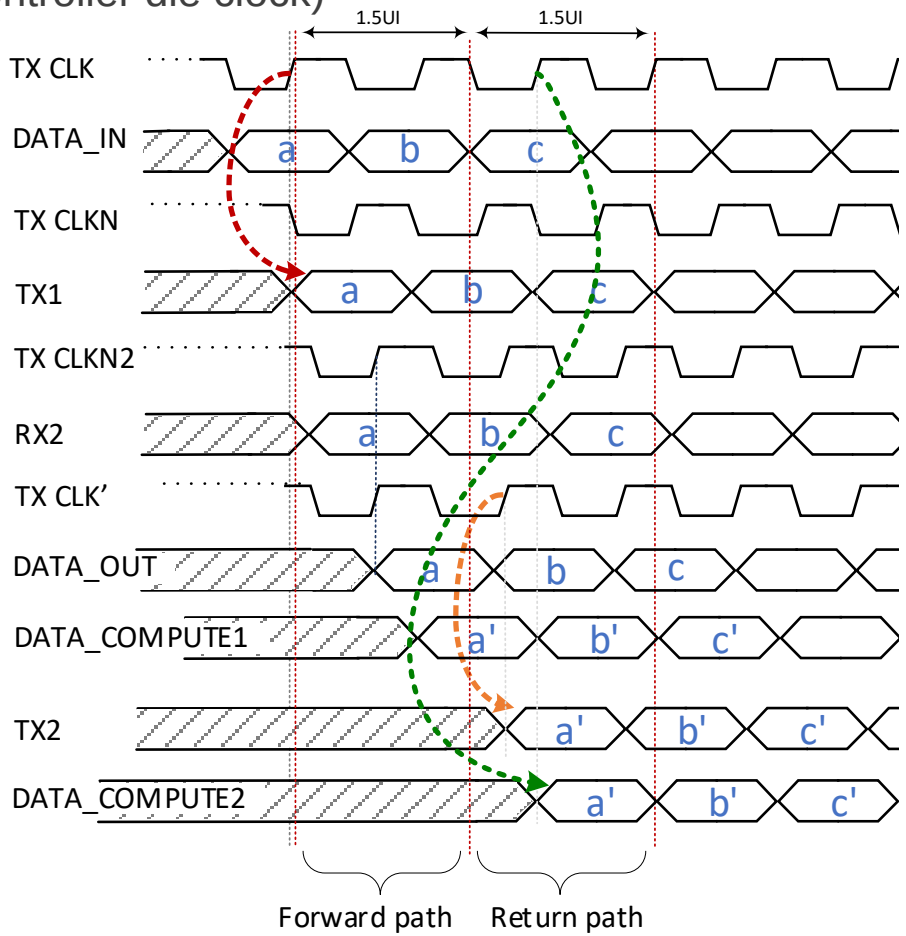
- Generator-based analog/mixed-signal + behavioral RTL
- Flexible across process
- Agnostic across D2D integration fabric (e.g., EMIB, interposer, …)
- Opensource

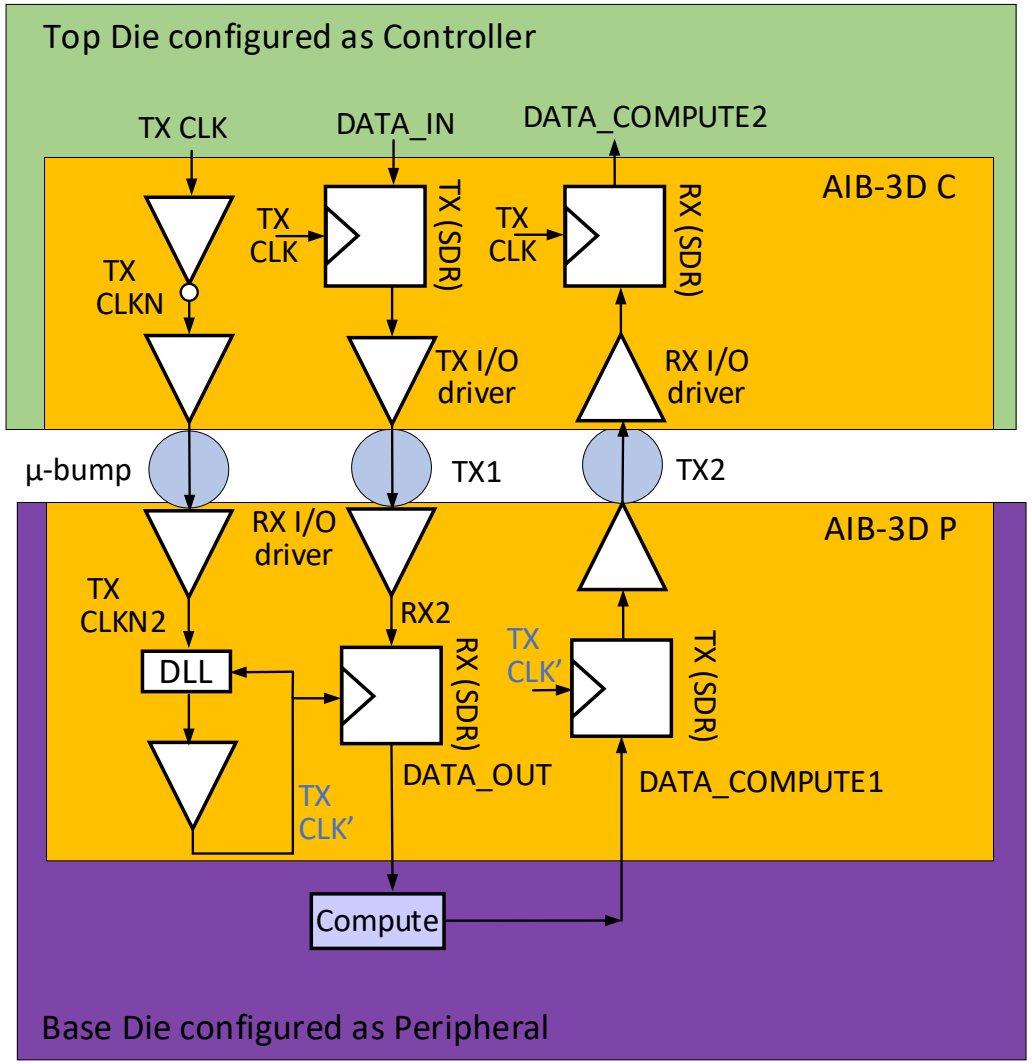
Desire similar features as 2D interface

CHIPS ALLIANCE

AIB-3D Clocking

- Leverage AIB2.0 Base for lighter weight implementations that require minimal circuitry
- Single data rate (SDR) signaling where new data is transferred on one edge of the clock
- Input and output signals are registered (additional latency)
- Array has a single clock that is forwarded with TX data
- On return path, there is no forwarded clock (captured with controller die clock)

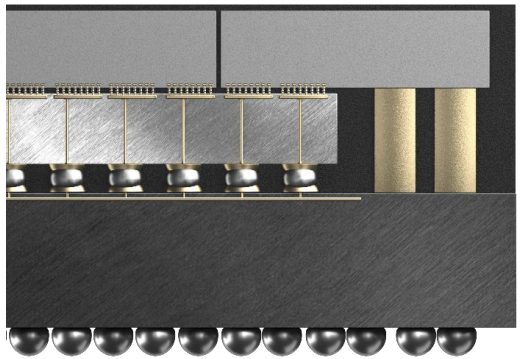




AIB-3D: Non-Intrusive Power Delivery

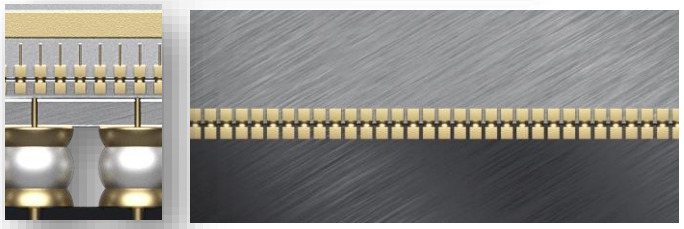
- Power delivery and thermal management is a challenge in 3D ICs
- Instead of forcing chiplets to incorporate power delivery to top die through the base chiplet (i.e., requires top die and base die co-design) leverage Foveros Omni [2] and Foveros Direct [3] to deliver power without compromising integrity of base die.

Foveros Omni [2]

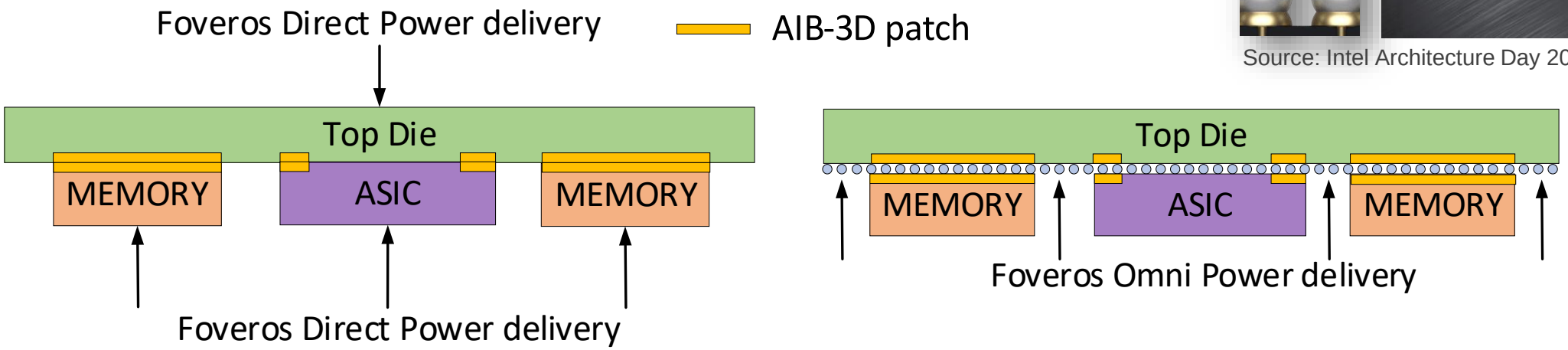


Source: Intel Architecture Day 2020

Foveros Direct [3]



Source: Intel Architecture Day 2020



Ref. [2]: A. A. Elsherbini et al., IEEE IEDM 2019
Ref. [3]: A. A. Elsherbini et al., IEEE ECTC 2021

AIB-3D: Summary

- Increasing use of 3D integration outside of just stacked memories
- Chiplet ecosystem in vertical direction requires standard interfaces
- AIB 1.0 and AIB 2.0 are opensource D2D PHY interfaces that enable high-performance interconnect → basis for 3D opensource standard
- AIB-3D:
  - Over-die and shoreline support
  - Scalable to support µ-bump / TSV scaling
  - Process agnostic, opensource
  - Support logic-on-logic and logic-on-memory
  - Options: orientation, modular tiling of "*AIB-3D patch*", configurable channels
  - Early specification in progress

CHIPS ALLIANCE