3D Die-to-Die Interconnect Testing Challenges and the Need for an IEEE Standard

Sreejit Chakravarty, IEEE Fellow
Principal Engineer, Intel Corporation

Introduction. The increasing trend to integrate and package multiple dice into one SoC has been driven by yield, i.e. cost, considerations, need to integrate proprietary functionality from different sources into the SoC, etc. Semiconductor manufacturers responded by developing innovative interconnect and packaging technologies, including 2.5D and 3D variants. In 2.5D, dice are placed next to each other on a substrate and interconnected using special IOs.
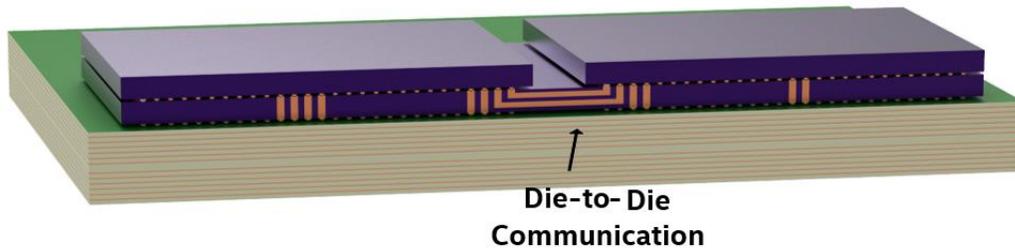

Figure 1. Cartoon of a 3D Stacked SoC showing the use of Die-to-Die communication

In 3D stacking, dice are placed on top of each other. Figure 1 shows an example SoC where two dice are stacked on top of a third die. Signals run between each of the two top die and the bottom die as well as between the two-top dice. To carry these signals, the collection of dice is interconnected, as shown, using 3D die-to-die interconnect. In this article we use the term die-to-die interconnects to refer to 3D die-to-die interconnect and focus on the test and repair of such interconnects.

Logical Model of the Die-to-Die Interconnect

The definition of the test and repair problem for die-to-die interconnects is based on the notion of a **cluster.**
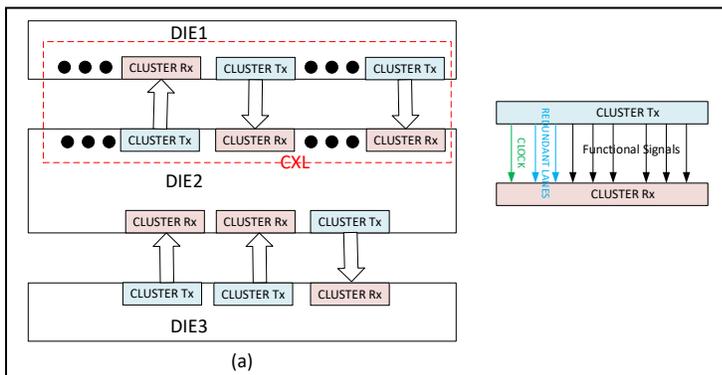

Figure 2. Clustering

Figure 2(a) show a 3-die stack. Standard buses run between these dice. A CXL bus connecting Die1 and Die2 is shown. CXL consists of signals from Die1 to Die 2 and vice-versa. Each signal uses a physical

interconnect between Die1 and Die2. Designers splits such standard functional buses into clusters. As shown in Figure 2(b), a cluster of size N, carrying N signals, consists of: a transmit end (Tx); a receive end (Rx); N die-to-die interconnects; 1 or more redundant die-to-die interconnects; and a forwarded clock going from Tx to Rx end. Redundant interconnects are used for cluster repair. The clock is used for synchronous data transfer between the two dice.

There are many variants of the individual die-to-die interconnects. This includes, but is not limited to: (i) face-to-face bonding, using solder bumps; (ii) face-to-face binding, using direct cu-cu connection; (iii) face-to-back bonding, which includes TSVs, etc. These variations impact the physical and electrical characteristics like area, density, performance, power-dissipation, etc. The logical abstraction of such connections is shown in Figure 3.
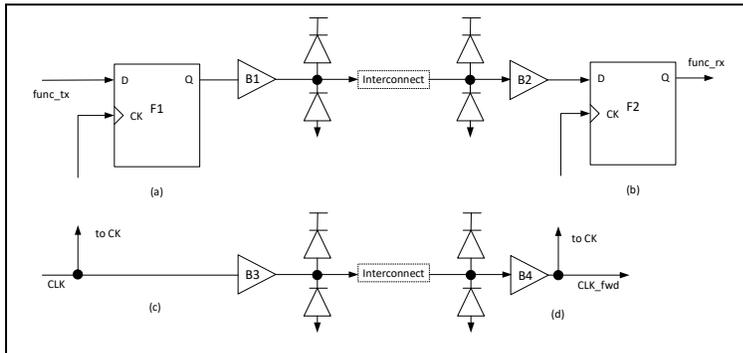

Figure 3. Logical abstraction of the Die-to-Die Interconnect

Figure 3(a) shows flip-flop (FF) F1 and driver B1 at Tx end, which are part of the Tx die. Figure 3(b) shows FF F2 and receiver B2 at Rx end, which are part of the Rx die. The "Interconnect" box between Figure 3(a) and Figure 3(b) abstracts the solder bump, cu-cu bonding, TSVs etc. used for connecting B1 to B2. Figure 3(a) and Figure 3(b) assumes synchronous data transfer, where a common clock is used transfer data from F1 to F2. Signals, like CLK in Figure 3(c) and Figure 3(d), do not have a common clock and are transferred asynchronously. In addition to clocks, other examples of other asynchronous signals are resets, powergood, etc. Synchronous connections and asynchronous connections are clustered separately resulting in: (i) synchronous clusters; and (ii) asynchronous clusters. Henceforth, cluster refers to either of these two cluster types.

Die-to-Die Interconnect Test and Repair Feature Description

Solutions to the following problems are required to implement the die-to-die test and repair feature.

- **No touch testing using standard DFT interface**. The footprint of die-to-die interconnect is very small. The footprint, and spacing, between such interconnects started with 50 μm and trending to be less than 10 μm. The number of interconnects vary from 1,000, for small designs, to 100,000 for larger designs. Physical probing of such densely packed signals, without sacrificing density, is not feasible. Therefore, a no touch test, repair, and debug method, using standards like IEEE 1687 to access the DFT infrastructure for these interconnects, is required.
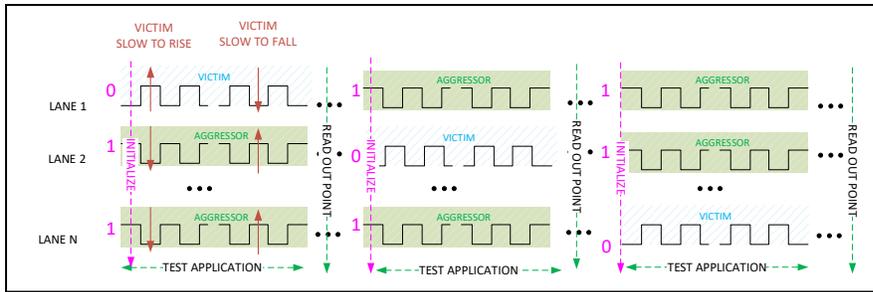
Figure 4. Victim-Aggressor Tests (VATs) for die-to-die interconnects.

- **Tests.** Specialized tests for such interconnects are needed. They differ from scan tests used for the logic area on the die. Test for logic target random defects which are modelled using single stuck-at or single transition faults. Tests for stuck-at faults are run at slow speed whereas tests for transition faults are run at-speed. Such tests are typically created using an ATPG tool and applied using scan DFT.

  A better test for die-to-die interconnects is the Victim-Aggressor Test, abbreviated as VAT, shown in Figure 4. For a cluster of size N, VAT consists of N phases: Phase 1 targets lane 1 where lane 1 is the victim; Phase 2 targets lane 2 where lane 2 is the victim; and so on. In Phase 1, where Lane 1 is the victim, signals on lane 1 transitions in one direction whereas signals on all other lanes, referred to as aggressors, transitions in the opposite direction. The result of the test is read out after each phase. This is repeated for other lanes in subsequent phase. In addition to detecting single stuck-at and transition faults on individual lanes VAT detects failures caused by capacitive or resistive coupling between the victim lane and aggressor lanes. Detecting coupling induced failures is important since: (i) these signals run in very close proximity; and (ii) the signal frequency is extremely high, running into several GHz.

  General purpose IO (GPIO) tests use pseudo-random tests, often referred to as pseudo-random bit stream (PRBS) tests, to accomplish the same goal. Simulation studies performed, using our latest interconnect models, show VATs to be at least as effective as PRBS with considerably reduced clock cycles. Thus, use of VAT reduces test time and test cost. Apart from test quality, another advantage of using VAT is that it enables implementation of an autonomous on-die diagnosis and repair functionality. Implementing on-die diagnosis and repair functionality using PRBS is extremely costly and therefore not feasible.

- **Analog Tests and Analog Observability.** GPIO test infrastructure includes extensive support for: (i) no-touch leakage (NTL) tests; and (ii) observing analog waveforms at SoC IOs. No such infrastructure is required for IOs driving die-to-die interconnects for the following reasons. NTL assumes that the difference in leakage current distribution between good and defective IO is such that an NTL threshold can separate good from defective parts. For IOs used for die-to-die interconnects, setting an NTL threshold is not feasible since there is a significant overlap of leakage current distribution resulting from defects and process variation. As Figure 3 shows, IOs used for die-to-die interconnects contain negligible analog components. Therefore, VAT tests run at slow and at-speed are adequate.

- **Debug and Failure Analysis.** After a cluster fails, root-causing the failure consists of two steps: (i) identifying the failing lane; and (ii) identifying if the defect is in the Tx die, the Rx die or the interconnect. In addition, fab houses use additional non-destructive probing techniques, like thermal sensing, to locate failures. Support for steps (i), (ii) and non-destructive probing is required.

- **Repair.** Given the large number of die-to-die interconnect and the complexity of implementing them defective interconnects are a fact of life. Significant yield improvements are achieved, in current interconnect generation, using die-to-die repair. An on-die autonomous hardware solution which tests the cluster, identifies the faulty interconnects, repairs, and retests the cluster is required. The repair solution must be fused into the die for subsequent use. An effective repair solution must comprehend the kind of defects, e.g. point, bridging or clustering as well as TSV failure rates.
- **System level test and repair.** Analogous to power-on self-test (POST) for memories, system houses require system level trigger mechanism that tests, and possibly repairs, die-to-die interconnect. POST addresses infant mortality issues, and later in the product life cycle, improves system availability.
- **Security.** Any feature providing access to die internals is a security risk. A requirement for any die-to-die test and repair solution is to mitigate such security risk.
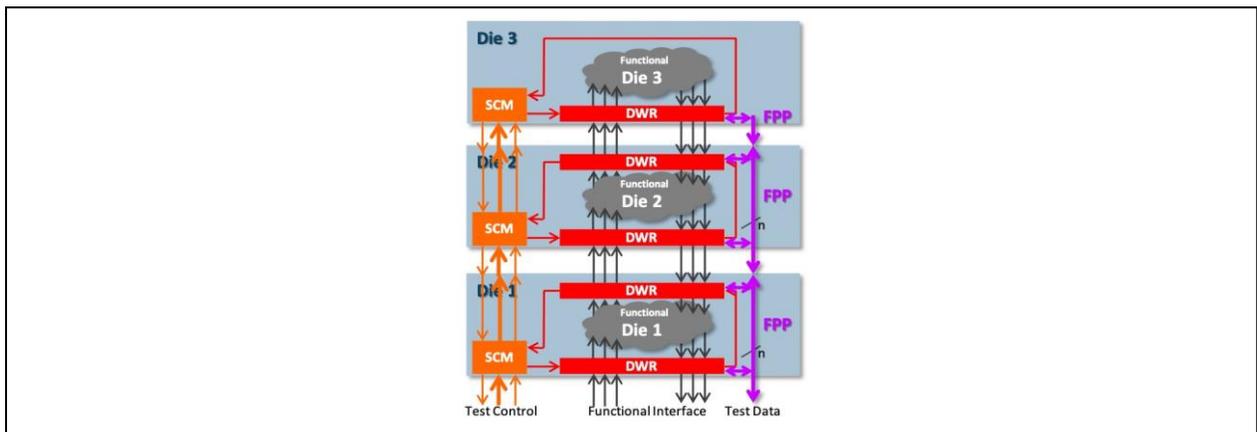

Figure 5. IEEE P1838 3D DFT Standard

- **BurnIn and Reliability.** Like all new process, especially with shrinking geometries, it is important to understand the reliability of these die-to-die interconnects. This must be translated into burn-in requirements.

Need for a new standard for die-to-die interconnect testing

There is an increased need to integrate die from different die-vendors. To facilitate proper test and repair of the packaged parts the above list of test and repair problems must be adequately addressed. An IEEE standard is required to address this problem.

The IEEE P1838 standard, a block diagram of which is shown in Figure 5, has standardized the DFT access mechanism for 3D stacked die. A standard has been developed for test control, using standardized SCMs, and test data, using standardized FPP. However, for die-to-die interconnect, it assumes a die-wrapper-register (DWR) and assumes this die wrapper to be part of the scan DFT infrastructure of individual die. This is not adequate since scan is not a good approach to solve the list of problems listed above. A wholistic view of the die-to-die test and repair solution is required and an autonomous, hardware test mechanism for die-to-die interconnects is needed. To facilitate effective die integration an IEEE standard that comprehends solutions to the above problems while understanding the various interconnect types is required.