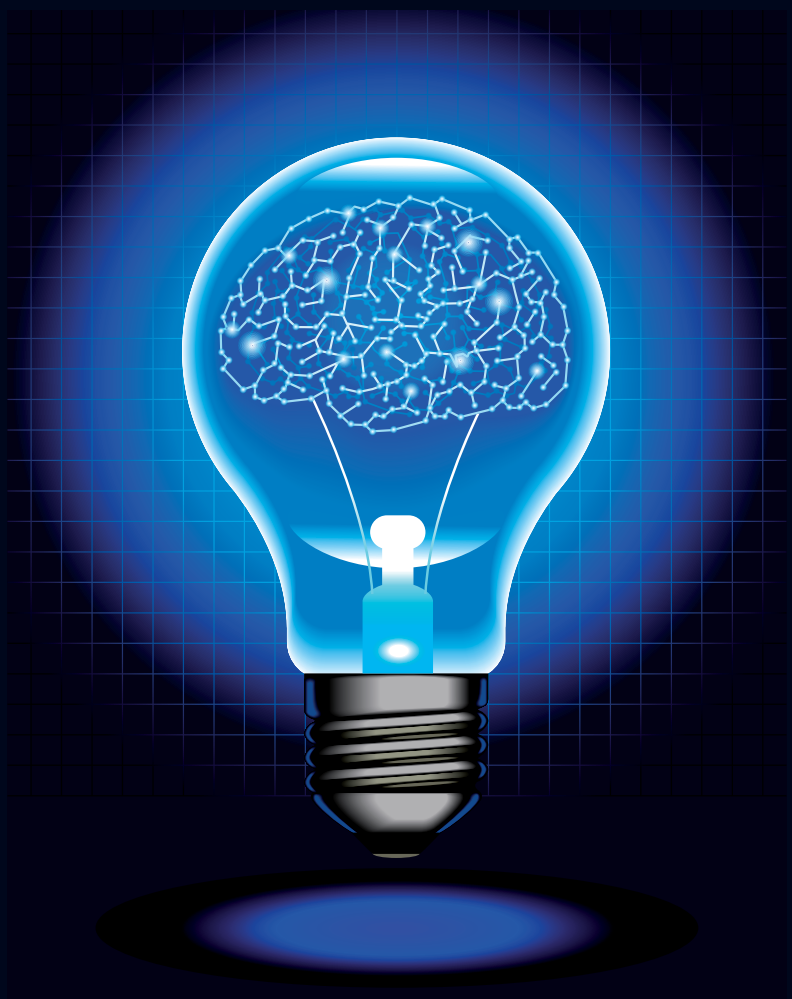
HiPEAC Technology Transfer Award winners

The brains driving new business forward

What do a cost model for integrated circuits, low-power network-on-chip technology and wearable body sensors have in common? They were all recipients of a HiPEAC Technology Transfer Award in 2015. Here, winners explain how they took their innovation out of the lab and into the marketplace.



OPTIMIZING 3D STACKED INTEGRATED CIRCUIT TEST FLOWS WITH 3D-COSTAR

Said Hamdioui, Delft University of Technology

Selecting an effective test strategy for a 3D stacked integrated circuit (3D-SIC) is crucial for overall cost optimization. In addition, diverse products and applications require different quality levels, resulting in different test flows; these flows may require different design-for-test (DfT) features, which need to be incorporated in the various dies early on in the design process. An appropriate cost model to optimize test flows with their associated DfT, while taking into account yields and die production costs, is therefore of great importance.

Delft University of Technology and IMEC have developed a theoretical foundation for 3D-SIC cost modelling, compiled into a tool called 3D-COSTAR. First announced in 2013 (http://bit.ly/press_release-3dcostar), the tool was later made available for public use on the Delft University website: http://bit.ly/3Dcostar.

Production flow from start to finish

3D-COSTAR is the first tool to address test flow and test cost optimization for 3D-SICs. The uniqueness of the tool is due to the fact that its inputs parameters cover the entire 2.5D-/3D-SIC production flow (design, manufacturing, test, packaging and logistics), and that it is aware of the stack build-up (2.5D versus 3D, multiple towers; face-to-face or face-to-back) and stacking process (die-to-die, die-to-wafer or wafer-to-wafer). The tool produces three key analysis parameters: 1) product quality, expressed as defect level (test escape rate) in DPPM (defective parts per million); 2) overall stack cost; and 3) breakdown per cost type.

3D-COSTAR allows precise, efficient trade-offs to be made at an early design stage. For instance, optimizing the test flows based on yield and cost parameters of an individual product is a

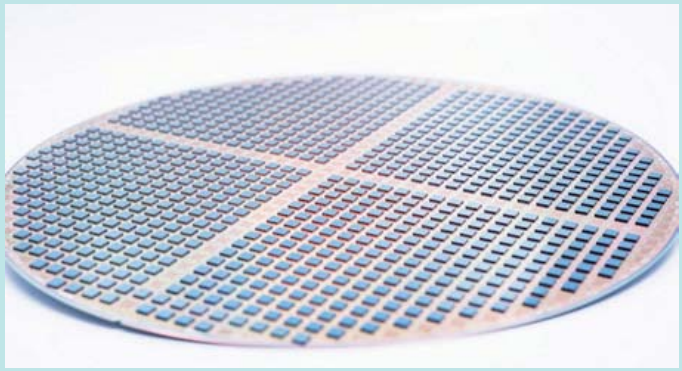


Image copyright: IMEC

complex optimization problem and strongly application dependent. Different test flows, executed after manufacturing, may require different design-for-test features, which need to be incorporated in the various dies during their early design stages.

The industrial importance of 3D-COSTAR is demonstrated by analysing trade-offs of different complex optimization test problems in terms of test, quality and cost. For example, the impact of pre-bond testing of active dies using either dedicated probe-pads or directly probing on large-array fine-pitch micro-bumps has been demonstrated in collaboration with IMEC and Cascade Microtech. These impressive results were presented at the IEEE International Test Conference in September 2014 in Seattle, USA. This work was also nominated by Semiconductor Equipment and Materials International (SEMI), a global industry association, as one of the four most influential Automatic Test Equipment (ATE) papers of the year in 2014.

Taking 3D-COSTAR to industry

IMEC has been using the tool in collaboration with some of leading companies in 3D-SIC in order to explore and analyse complex trade-offs in 3D test flows, in terms of both cost and DPPM. For example, together with IMEC, Cascade Microtech has recently demonstrated the feasibility of direct probing large-array fine-pitch micro-bumps to avoid the usage of dedicated pre-bond pads. Analysis with 3D-COSTAR clearly showed up to 50% overall cost benefit of doing micro-bump probing using an advanced probe cell such as was demonstrated with Pyramid Probe® RBI technology on Cascade Microtech's CM300 probe station.

The target market for 3D-COSTAR is all companies involved in 3D- and 2.5D-stacked ICs – that includes just about every semiconductor company. We are reaching out to them via papers and presentations at international conferences and trade-shows, via IMEC's strong network of industrial partners, and by making a version of the tool publicly available via a website of TU Delft.

In my view, to transfer research results you need a deep understanding of the real needs of industry and collaborate with industry partners from the very start of the development process. This allows the solution to be refined so that it meets industry needs.

NETWORK-ON-CHIP FOR THINK SILICON'S ULTRA-LOW-POWER GPU

Giorgios Dimitrakopoulos, Democritus University of Thrace (DUTH)

Based near Patras in Greece, Think Silicon S.A. develops high-performance, ultra-low power graphics intellectual property (IP) for embedded and mobile applications. The company currently focuses on devices for the wearables and Internet of Things (IoT) market, where the main challenge is extending battery life.

Think Silicon required a power-efficient, scalable Network on Chip (NoC) technology that would integrate its multicore NEMA 2D and 3D Graphics Processing Units (GPUs) and display processors. The company was interested in a scalable interconnect solution that would allow them to seamlessly customize their available multicore configuration in accordance with customer requirements, without paying for additional redesigns on a customer-by-customer basis. This has a direct positive impact to time to market, a key factor in decision making particularly for SMEs targeting the highly evolving IoT market.

Technical advantages over the competition

The NoC we created, TSNoC, allows NEMA 3D GPU IP to be automatically integrated to a host System-on-Chip using configurable AMBA AXI-based interfaces. TSNoC was tailored to the increased performance requirements of GPU-specific memory traffic patterns without exceeding the ultra-low power budget or the tight space constraints of wearable or IoT devices. To tackle both challenges, TSNoC is optimized for the traffic and memory patterns of the NEMA GPU, and leverages the cores' multithreading capabilities to offer fast data delivery.

A proprietary arbitration policy and load-balancing flow-control protocol allow fair access to the memory controller across all threads, while any inter-core traffic can be efficiently isolated by the core-to-memory traffic. TSNoC can support an arbitrary number of memory ports to handle GPU memory traffic regardless of address mapping across ports and across threads. In this way, software development and thread scheduling inside the cores is simplified, thus minimizing hardware complexity and power consumption.

TSNoC is also associated with a rich verification framework, consisting of constrained random transaction generators acting as AXI Master and Slaves and proper transaction delivery verifica-