Cost-Effective Design of Scalable High-Performance Systems Using Active and Passive Interposers

Dylan Stow, Yuan Xie
Electrical and Computer Engineering
University of California, Santa Barbara
Santa Barbara, California
{dstow, yuanxie}@ece.ucsb.edu

Taniya Siddiqua, Gabriel H. Loh
AMD Research
Advanced Micro Devices, Inc.
Bellevue, Washington
{taniya.siddiqua, gabriel.loh}@amd.com

Abstract*—Cutting-edge high-performance systems demand larger and denser processors, but future lithographic nodes are expected to introduce higher manufacturing costs and yield challenges. Die-level integration technologies like passive interposer-based 2.5D have demonstrated the potential for cost reductions through die partitioning and yield improvement, but system performance and scalability may be impacted. Alternatively, active interposer technology, the intersection of 3D and 2.5D methodologies, can provide higher-performance interconnect networks to integrate chiplets, but the active interposer die is itself subject to cost and yield concerns. In this work, we perform a cost and performance comparison between traditional monolithic 2D SoCs, 2.5D passive interposers, and 2.5D/3D active interposers to demonstrate the trade-offs between the interposer types for current and future high-performance systems. This work introduces a multi-die core-binning cost model to demonstrate the yield improvements from interposer-based die partitioning of large multi-core processors. The relative cost and performance scaling trade-offs of passive and active interposer dies are then compared for the target systems, demonstrating that both methodologies can indeed provide cost-effective integration for different system requirements. Finally, this work demonstrates how the extra "prepaid" silicon area of the interposers can be leveraged for fault tolerance to improve yield and cost-effectiveness. In summary, this work concludes that both active and passive interposers can cost-effectively improve the functional and parametric yield of high-performance systems, together providing a cost versus performance space to meet a range of design requirements.

I. Introduction

As outlined in the ITRS 2.0 roadmap [2] [4], the datacenter and microserver markets demand increasingly performant and localized processing, with a roughly 3× increase in available memory and 4× increase in the number of processor cores per socket and rack unit, respectively, over the next ten years. Similarly, the push for high-performance exascale supercomputing will likely require complex heterogeneous SoCs with many cores and integrated memory to provide sufficient bandwidth and data localization to meet efficiency requirements [20]. Modern manycore server processors, such as the 32-core AMD "Epyc" processor, demonstrate that the industry is indeed moving in these directions to meet datacenter and microserver demands.

Unfortunately, the ability to meet these demands with conventional process scaling is becoming increasingly difficult and expensive. The Moore's Law target cadence is already slipping, with almost all foundries no longer able to meet the desired transistor scaling rates in the most recent process nodes [12] and future process roadmaps slowing for each new node. Increased process complexity has led to more expensive fabrication and longer manufacturing cycle times [13], and as transistor cost reduction slows, yield and endurance challenges grow, and cost per area increases [24] [28], it becomes increasingly costly to meet the market requirements for denser, larger integrated circuits.

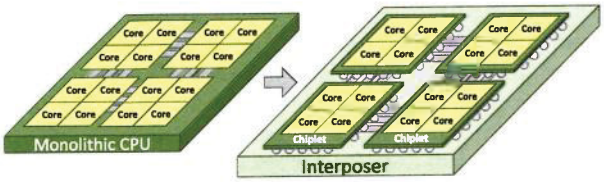


Fig. 1. Transition from monolithic manycore CPU to interposer-based 2.5D system with multiple chiplets

As shown in Figure 1, an alternative solution to traditional monolithic SoC integration is the usage of die-level integration methods like Through Silicon Via (TSV)-based 3D and interposer-based 2.5D methodologies. Manufacturing yield can be improved by partitioning the SoC into multiple chiplets, ideally with identical modular structure to reduce design and mask cost, and by bonding these chiplets through high-yield, high-bandwidth, chip-to-chip interconnects. 3D integration has long been studied as a solution to improve yield and performance, but die stacking requires significant EDA changes and leads to thermal density challenges. Interposer-based 2.5D integration, however, has already come to market for several high-end devices, including the AMD Radeon R9 GPUs with High Bandwidth Memory integration for improved performance, efficiency, and footprint [15] and the Virtex-7 FPGA from Xilinx [19] with multiple FPGA slices and heterogeneous transceiver chiplets for improved yield, configurability, and performance. However, the usage of interposers has so far been limited to these cases, while the wider high-performance market could stand to benefit from interposer adoption. In recent analysis of a cost-driven design methodology, both 2.5D and 3D designs were shown to have lower post-yield manufacturing costs than 2D SoCs for midsize and large systems [22], but only 2.5D designs were cost-effective for high-power designs, while 3D suffered from increased packaging and cooling costs when thermal management was considered [23].

978-1-5386-3093-8/17/$31.00 ©2017 IEEE

NVIDIA 1020

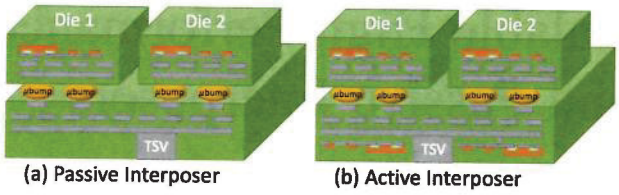


Fig. 2. Illustrative two-chiplet system, integrated with microbumps using (a) Passive interposer with only passive interconnect and TSVs, and (b) Active interposer with active CMOS logic.

Although interposers can be utilized for partitioning and integration, the metal-only nature of current **passive interposers** potentially limits their ability to provide sufficient bandwidth and latency for new high-performance systems. **Active interposers** [14] are an emerging combination of 2.5D and 3D integration that balances the simplified design methodology and thermal management of passive 2.5D but leverages standard CMOS processes to integrate active transistor devices into the interposer for faster repeated interconnect and flexible Network-on-Chip (NoC) for better chiplet connectivity [8]. Active interposers have been demonstrated to improve signaling and efficiency over passive interposer [10], [11], and functional samples with active NoC have recently been fabricated [26].

The transition from a passive to active interposer increases the interposer cost overhead due to additional process complexity, and the active interposer itself could become a large, low-yield die that increases system cost. To date, no active interposers have been adopted in commercial designs due to these cost concerns. As such, all recent active interposer work has focused on "minimally active" interposers [9], [26] with only a small percentage of the available area utilized to minimize yield losses. Some work has gone as far as simplifying the transistors to minimize the number of extra process steps, at the expense of transistor functionality [25]. Yet in all of these minimally active designs, a large and costly active CMOS die is being produced and paid for, but little effective area is being utilized.

This work explores the benefits and trade-offs of active and passive interposer-based design for high-performance systems. First, the yield and performance benefits of interposer-enabled die partitioning are demonstrated in Section II through the use of a novel core-binning 2.5D cost model. Following this justification for interposer-enabled partitioning, Sections III and IV provide guidance on interposer technology selection through analysis of active and passive interposers on the metrics of performance scalability and cost overheads. Further, fault-tolerant methods are proposed to reduce active interposer cost overhead without increasing total system footprint. This work conflicts with prior assumptions about active interposer cost-effectiveness and demonstrates the feasibility, with proper technology selection, of both active and passive interposer design methodologies to provide cost reductions and high bandwidth integration for a broad range of high-performance systems.

II. The Case for Interposers: Yield and Binning Improvement from Die Partitioning

Modern and future performance-targeted systems will span the wide market range from desktop CPUs and GPUs used for virtual reality and workstation applications, to exascale processors for the most demanding scientific and big data computations. Unlike the mobile and IoT markets, these high-end systems have significantly larger die sizes and thus more difficult yield challenges. For consumer devices, such as an eight-core desktop and workstation processor, manufacturability translates to improved performance per dollar. For manycore server processors [1] or future exascale processors [20], improved yield and reduced manufacturing costs allow for lower total cost of ownership and wider market share. These cost reductions, pushed down to the consumers and warehouse-scale providers, allow for the proliferation of higher-performance processing, thus expanding the range of achievable software solutions across the field.

In this section, we demonstrate how interposer-enabled die partitioning can result in significant manufacturability improvements and cost and functionality benefits across the range of performance targeting circuits, motivating the transition away from monolithic SoC integration. First, a manufacturing yield and cost model is presented for interposers and 2.5D systems. To improve the model accuracy for large-area circuits, a novel core-binning defect model and a chiplet matching strategy are developed. The application of these models on two case studies demonstrates how interposer-based partitioning can greatly improve yield and increase the number of high-margin, high-performance fully-enabled chips, especially if future processes exhibit yield challenges.

A. Manufacturing Cost Model for SoC and Interposer Systems

The cost of a single semiconductor die can be estimated by using only the die area and the process technology. The choice of the process technology has a major impact on the die cost, determining the cost per wafer C_w and density of critical defects D_0. Performance-targeted circuits historically adopt the most recently available process technologies to leverage the latest improvements in transistor density and speed, although it remains to be seen how future technologies will scale in cost, yield, and reliability. The defect density D_0 of a new process is initially high, but it decreases, generally by 2-5× for historical technologies [6], over several years as the process matures. Using the negative binomial yield model [21], the yield of an individual die can be calculated with critical area A as:

$$Y_{die} = \left(1 + \frac{A * D_0}{\alpha}\right)^{-\alpha} \quad (1)$$

where α is a process dependent clustering parameter, frequently between 1 (high defect clustering) and 3 (moderate defect clustering)[1]. For logic-dominated dies, the critical area A is commonly assigned to be the total area of the integrated

[1]Poisson yield, with uniform defect distribution, is overly pessimistic for large dies [6], but can be approximated with $\alpha \geq 10$.

circuit. With the die yield, die area, and wafer diameter ϕ_{wafer}, the number of dies per wafer N_{die} is found with:

$$N_{die} = \frac{\pi \times (\phi_{wafer}/2)^2}{A_{die}} - \frac{\pi \times \phi_{wafer}}{\sqrt{2 \times A_{die}}} \quad (2)$$

The manufacturing cost per die is then calculated as:

$$C_{die} = (\frac{C_{wafer}}{N_{die}})/Y_{die} \quad (3)$$

where C_{wafer} is the process-dependent wafer cost.

These three equations are sufficient for modeling the manufacturing cost of a single 2D semiconductor die, but a 2.5D interposer-based system introduces additional cost overheads. Unlike stacked 3D integration, the primary active dies in a 2.5D system do not require thinning or through-silicon via (TSV) creation. The dies are bonded to the interposer using face-to-face (metal-to-metal) bonding through microbumps, copper pillars, or micropads. This bonding process does however introduce an extra process complexity that translates into fabrication cost and a potential failure point that can influence yield. Thankfully, bonding assembly yields have been consistently demonstrated at greater than 99% success rates [11], [15]. Of course, the interposer-based system must also include the cost of the interposer itself, which can again be calculated like a standard die using Equations 1-3, with adjustment only to the wafer cost C_{wafer}. A passive interposer only has TSVs (to connect to the substrate) and several layers of metal interconnect, so the wafer cost is significantly lower than a comparable CMOS process technology (explored in detail later in Section IV-B). An active interposer would be fabricated by using an existing CMOS process technology and by then adding TSVs, resulting in a higher cost per die than a passive interposer given the same size and yield. The total manufacturing cost of a 2.5D system with n chiplets and one interposer is calculated with:

$$C_{2.5D} = \frac{\frac{C_{int}}{y_{int}} + \sum_{i=1}^{n}(\frac{C_i}{y_i} + C_{bond_i})}{Y_{bond}^{n-1}} \quad (4)$$

where C_{int} and y_{int} are the interposer silicon cost and yield. In this work, we assume that Known Good Die (KGD) testing is performed on each chiplet before bonding to the interposer, which is necessary for improving system yield and reducing manufacturing cost [14].

Nonrecurring costs like engineering effort for design and verification, or production of mask sets, can also contribute to total cost per die, especially when volumes are low. Because the high-performance systems under examination already require the largest design effort and are appropriately marketed in large volumes, we assume that any nonrecurring costs are sufficiently amortized across volume or are minimally changed between integration approaches.

B. A Core-Binning Yield Model for Modular Circuits

The yield model in Equation 1, although commonly used in prior work, are not representative of the fabrication of large-area integrated circuits. With chip sizes that can approach the reticle limit, the yield for a defect-free die can be very low, even for mature process nodes. For example, according to Equation 1 with $\alpha = 3$, a 600 mm^2 GPU die in a mature process node with defect density $D_0 = 0.2\ cm^{-2}$ [17] would have a die yield (before parametric variation) of only 36%. For an emerging process with $D_0 = 0.5\ cm^{-2}$, yield is only 12.5%! In order to improve revenue and produce more functional parts, leading manufacturers of CPUs, GPUs, and other high-performance circuits rely on **binning** at the core unit level. If a defect is present in a modular core, the impacted segment of the die is disabled and the chip is sold with reduced functionality at a lower price.[2]

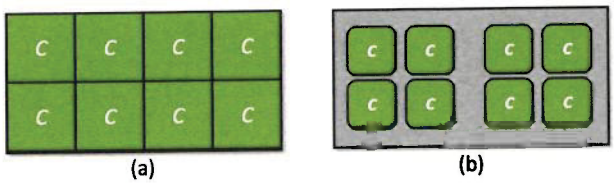


Fig. 3. (a) An eight-core die in which 100% of the die can be flexibly disabled for binning. (b) A representative system where only the cores, which make up 50% of the area, can be disabled for binning.

In order to model the distribution of defects between and within the dies, we utilize the derivation equation of the negative binomial yield model, shown below in Equation 5.

$$P_{defect} = \frac{\Gamma(d+\alpha)}{d! * \Gamma(\alpha)} * \frac{\beta^d}{(\beta+1)^{d+\alpha}} \quad (5)$$

The probability that a die has d defects is calculated using the gamma function $\Gamma(x)$ and constant β defined as:

$$\beta = \frac{D_0 A}{\alpha} \quad (6)$$

Within the relatively local area in a single die, it is assumed that defects are randomly distributed (Poisson) across the cores and uncore area. Multiple defects may fall into the same core, resulting in more functional cores after binning. The probability of a die with d defects and c binnable modular cores to have g good, functional cores is:

$$P_{good} = \frac{S(d, c-g)\binom{c}{c-g}(c-g)!}{c^d} \quad (7)$$

where $S(d, c-g)$ is the Stirling number of the second kind. Equation 7 assumes that the whole die is partitionable for binning. In real designs, non-modular uncore units like interconnect fabric and system management contribute significant die area and are not easily disabled. Figure 3 shows an eight-core processor with (a) fully partitionable die area (b) 50% binnable core area and 50% critical uncore area, representative of modern designs. Equation 7 can be expanded to account for non-modular critical area percentage η:

$$P_{good_\eta} = P_{good} * (1-\eta)^d \quad (8)$$

[2]Although mobile systems have grown in heterogeneous complexity, high-performance systems continue to scale along modular units, with benefits to design and software effort. For simplicity, the analysis here addresses the most common homogeneous systems, but it is similarly applicable to heterogeneous systems of sufficient modularity.

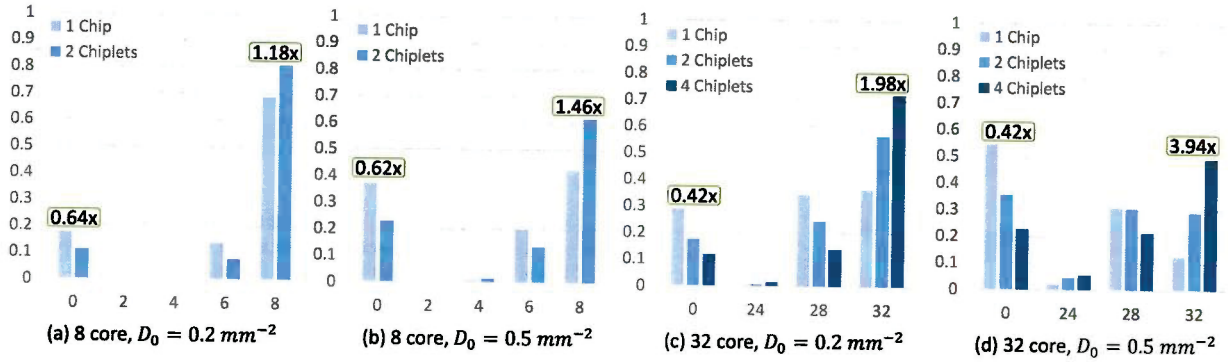


Fig. 4. Yield distribution of binned dies after manufacturing for each functional core count bin.

C. Core Binning and Cost Results for Eight-Core Processor

By taking the sum of products of Equations 5 and 8 across all defect counts, we can determine the yield distribution for each number of functional cores. We first apply our models to investigate a mainstream eight-core desktop/workstation consumer processor with $A = 200\ mm^2$, $\alpha = 3$, and $\eta = 0.5$, as shown in Figure 3b. Binning is performed at multiples of two cores, as in modern commercially available processors. For the two-chiplet design, a greedy matching process is used to produce as many fully-enabled processors as possible. A per-chiplet bond yield Y_{bond} = 99% [15] is included in the two-chiplet system yield distribution to reflect pessimistic integration losses. Binned yield distribution results are shown in Figure 4 (a) and (b) for two potential defect rates: a mature process with $D_0 = 0.2\ cm^{-2}$ and a cutting-edge process, or potentially a low yield future process, with $D_0 = 0.5\ cm^{-2}$. The yield improvement from chiplet partitioning and KGD testing translates to a reduction in unsalvageable chips and an increase in the number of fully enabled, high margin chips. At the defect rates for a mature process and an emerging process, the number of fully functional cores is estimated to increase by 1.18× and 1.46× and the number of failing systems decreases by 0.64× and 0.62×.

Speed Bin	2 core	4 core	6 core	8 core
Target	1	1.7	2.5	5
Slow	0.8	1.5	2	3.7

TABLE I
NORMALIZED PRICE PER CORE COUNT OF EXISTING CONSUMER PROCESSORS AT TWO SPEED BINS.

To measure the total utility of these improvements to yield and functionality, we can utilize the estimated price of equivalent commodity processors as a representative value metric. Table I lists normalized, approximate price ratios for each core count at two speed bins based on previously published consumer devices [3]. To model parametric yield, which can also be improved through die partitioning and known good die matching [9], a Gaussian frequency distribution is assumed for each core, with any cores with frequency below one standard deviation of the mean binned to "Slow" and average and faster cores binned to "Target." Under this simple parametric model, about half of the four-core chiplets will achieve the target speed, while only a quarter of the eight-core chips can meet the target. Through a combination of functional and parametric yield improvements, the utility value metric of the two-chiplet system is improved by 20.8% when $D_0 = 0.2\ cm^{-2}$ and by 41.4% when $D_0 = 0.5\ cm^{-2}$.

D. Core Binning and Cost Results for 32-Core Processor

While modest yield improvements are seen from chiplet partitioning for the consumer processor at mature defect densities, increasingly significant gains are seen for larger area circuits like server processors that exhibit greater yield challenges. Yield distributions for an example 32-core server processor with $A = 600\ mm^2$ are shown in Figure 4 (c) and (d) for the same $D_0 = 0.2\ cm^{-2}$ and $D_0 = 0.5\ cm^{-2}$, respectively. Die partitioning results in a 0.42× reduction in failing chips and a 1.98× improvement in the number of fully enabled chips for the mature process, and a 0.42× reduction in failures and a very sizable 3.94× improvement in full enablement for the emerging process.

III. Interposer Selection: Performance and Scalability

In the previous section, significant improvements in manufacturability are shown from the chiplet partitioning of large monolithic systems. This technique can be enabled by multiple emerging packaging technologies, but the requirements for high bandwidth, high efficiency, and low latency in performance-targeting systems are difficult to achieve with coarse-featured package-level integration techniques. The fine-featured die-level integration of passive or active interposers, however, is able to concurrently meet these performance goals. Within this interposer design space, circuit-level differences between active and passive interposers determine the feasible NoC architecture designs and resulting performance. In this section, we analyze these interposer NoC architectures in terms of scalability, area overhead, and link frequency in order to assist designers in the proper interposer technology selection to meet system requirements.

A. Active and Passive Interposer NoC Design

The interconnect-only nature of passive interposers, versus the embedded routers and low latency repeated wires of active interposers, leads to major differences in NoC design between

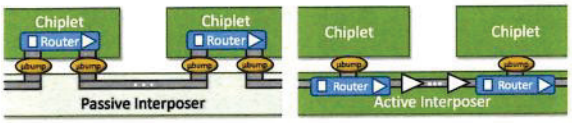


Fig. 5. NoC integration topology for passive and active interposer.

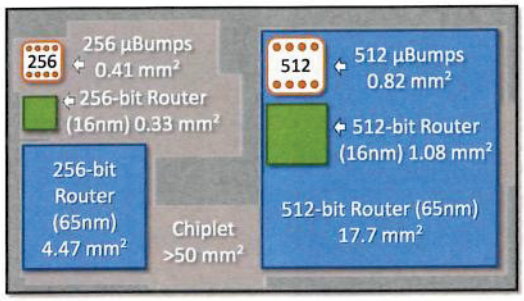


Fig. 6. Scale comparison of 40 μm pitch microbump arrays to 256-bit and 512-bit flit width routers in 16nm and 65nm technologies.

the two interposer types. For the passive interposer, all routers must be fabricated into the chiplet dies, contributing chiplet area overhead. Each network link is driven from the output channel through the microbumps into the passive interposer, where it travels along a long unbuffered interconnect link before again passing through a microbump to the receiving router input channel. With routers in the chiplets, all inter-chiplet NoC links, in all directions, must pass through these die-die connections, which often include electrostatic discharge (ESD) protection overheads [25]. The active interposer, however, only needs to add a single high bandwidth hop from a chiplet node to an on-interposer router. Within the active interposer, the flit can be passed between routers without the overhead of die-die microbump transmissions. Additionally, repeaters along the links can reduce interconnect transmission delay and increase the achievable network frequency. The increased design flexibility of the active interposers, with reduced constraints on microbump utilization and router placement, presents a wide range of network architecture opportunities to meet performance requirements [9], which for exascale systems may be multiple terabytes per second of memory bandwidth [20]. The network architecture differences between interposer types are demonstrated in Figure 5.

1) Router and Microbump Technology Scalability: One necessary design consideration for interposer-based NoC is the area scalability of the microbump arrays versus the area of the process technology-dependent routers. Modern microbump technology is standardizing on 40μm pitch, with potential reduction to 5 μm pitch in the future [18]. At current pitches, a 512-bit link spans an area of at least 0.82 mm^2 (not including any local microbump allocation for power or clocks), and a 256-bit link is half this area at 0.41 mm^2. A 5x5 router for a passive interposer will have 2 unidirectional links internal to the chiplet and 8 through-interposer links for the 4 cardinal directions, thus requiring 8 microbump arrays of the link width. This is still a reasonably small percentage of the peak available chiplet bandwidth (only 13% of even a small 50 mm^2 chiplet), but it could limit the number of routers per chiplet. Of more significant concern is the scalability of microbump pitch with router size. Using the *McPAT* modeling framework for quick and repeatable estimation, the areas for a 5x5 NoC router can be generated for a range of process technologies from 16nm to 65nm and beyond [16]. Figure 6 illustrates potential scaling issues for both active and passive interposers. For passive interposers, the area of a router in a modern 16nm process is slightly smaller than the area of a single microbump array of the same width, but because 8 unidirectional links are required between the chiplet and passive interposer, sufficient fan-out wiring must be added, further consuming chiplet resources. For an active interposer in an aging technology node like 65 nm, the router area can be an order of magnitude larger than a single microbump array. This facilitates the low-overhead communication between the interposer and chiplet, but it limits to the number of routers in the active interposer when older processes are selected.

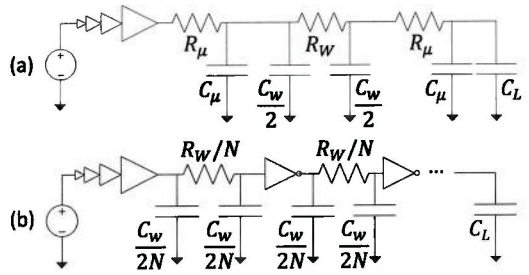


Fig. 7. NoC circuit differences between (a) passive and (b) active interposers.

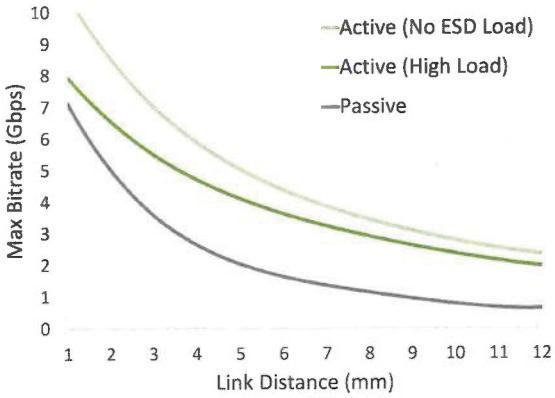


Fig. 8. Maximum bitrate versus link distance for the passive interposer, active interposer with same load capacitance as passive interposer, and active interposer without ESD load overhead.

2) Link Frequency in Active and Passive Interposers: As discussed in Section III-A, the lack of active devices in the passive interposer requires that routers are placed in the chiplet dies and that links must route through the die-die microbumps and across longer unrepeated interconnect.The circuit models for the different interposer types are shown in Figure 7 for the passive interposer link, with microbump RC, and for an active interposer link with N repeaters. To achieve high bandwidth and low latency routing, the active interposer has the advantage of lower RC (without die-die

connections) and reduced interconnect delay from repeaters. Further, the die-die interconnect needs ESD protection on the bumps to protect the circuit during manufacturing, resulting in additional capacitive load for each passive interposer link. To model the difference in link delay and maximum network frequency, the circuits were simulated using HSPICE using the 65nm PTM models for transistor and interconnect [29]. For each specified link distance, the drivers and repeaters were optimized to minimize link delay. Maximum bitrate results are shown in Figure 8 for interconnect settings with 350 nm wire width and spacing [18], 1.2 μm thickness, starting driver width of 2x, and maximum repeater width of 64x. To demonstrate sensitivity, two curves are shown for the active interposer: one with the same capacitive load as the passive interposer with ESD protection overhead (200 fF) and one with a lower load of 50 fF. The microbumps, with self capacitance of only 15 fF [18], introduce limited overhead compared to the lengthy interconnect. The repeaters, however, provide a significant advantage to the active interposer, which is able to achieve several times less delay than the passive interposer for the same link length. The active interposer can thus provide a greater range of NoC performance, with reduced latency links for higher network frequency or longer physical links at the same frequency.

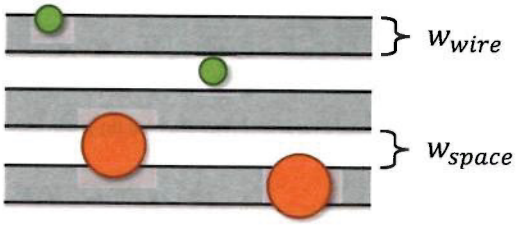


Fig. 9. Critical defects (large) cause shorts and cuts in the interconnect, while smaller defects are non-critical.

IV. Interposer Selection: Cost and Yield Overhead

As demonstrated in Section II, the partitioning of a large monolithic SoC into multiple chiplets can result in significant improvements to yield and functionality. Active and passive interposers are able to provide high bandwidth NoCs for chiplet reintegration to meet a range of performance requirements, as shown in Section III. Unfortunately, interposer fabrication and chiplet bonding add manufacturing cost overheads that may diminish the total system cost benefits. Additionally, although active interposers demonstrate lower link latency, higher bitrates, and more flexible NoC architectures, the extra process and design complexity versus passive interposers translates to further cost and yield overheads. In this section, we analyze the relative magnitudes of these overheads versus system cost improvements across a range of interposer technology choices. We find that active interposers are indeed consistently more expensive than passive interposers, but that with proper technology selection they are both cost-effective integration solutions for high-performance systems. Further, based on the presented yield and cost breakdowns, the "prepaid" vacant area of the interposer is leveraged for fault tolerance to reduce the active interposer cost overhead. In order to meet system requirements, system designers can leverage the analysis and techniques in this section to balance the cost versus performance trade-offs between active and passive interposers.

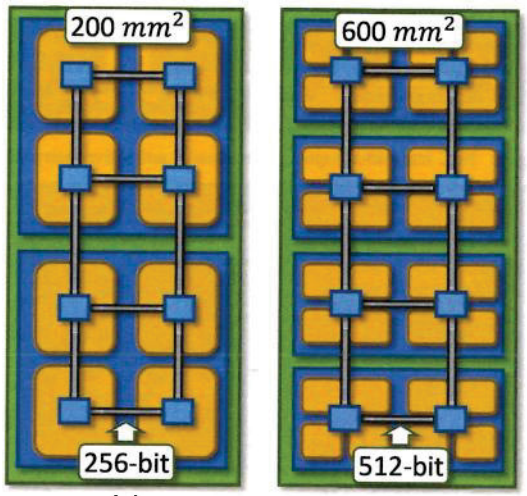


Fig. 10. Chiplet partitions and 2×4 NoC meshes for (a) eight-core and (b) 32-core systems.

A. Interconnect Yield Model for Interposers

Unlike most silicon circuits, a passive interposer is primarily metal interconnect, surrounded by vacant space. An active interposer is similar in design, but may also have sparse logic activity for routers and repeaters. The prior assumptions for Equation 1 for critical area and defect density are inaccurate for interconnect yield, since a wider route is instead more resilient to a small defects that would disrupt minimally sized features. As shown in Figure 9, failures occur as shorts between wires (in the same or adjacent layers) or as open cuts [5]. Large wires and spacings require larger sized defects to cause a failure, and historically densities for larger defect sizes drop quickly compared to the critical feature size [6]. Maximum density minimally sized wires will have lower yield, while wide or sparse interconnect improves yield. Based on prior studies [5], [6], we account for this increased resiliency by reducing the defect density from $D_0 = 0.2\ cm^{-2}$ to $0.05\ cm^{-2}$, in approximate proportion to the interconnect dimensions versus minimum dimension.

B. Interposer Cost and Yield Comparison

The interposer-enabled die-level integration introduces cost overheads into the manufacturing of each system, but the exact amount depends on multiple design decisions, including the interposer process technology (including active or passive), the number of chiplets that must be bonded, and the complexity of the interposer interconnect system that will influence yield. We study the systems shown in Figure 10: an eight-core processor partitioned into two chiplets and a 32-core system partitioned for yield improvement into four chiplets. To provide low

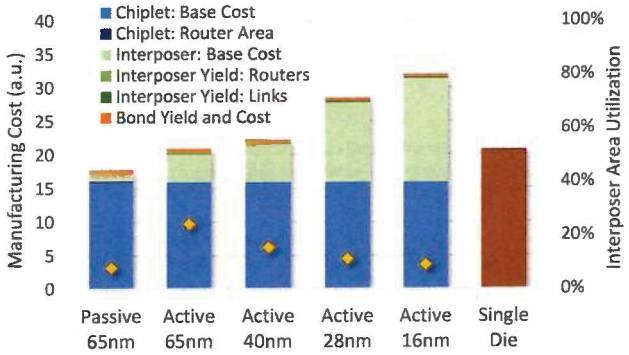


Fig. 11. Manufacturing cost breakdown of eight-core two-chiplet interposers systems versus 16nm monolithic die. Interposer utilization on secondary axis.

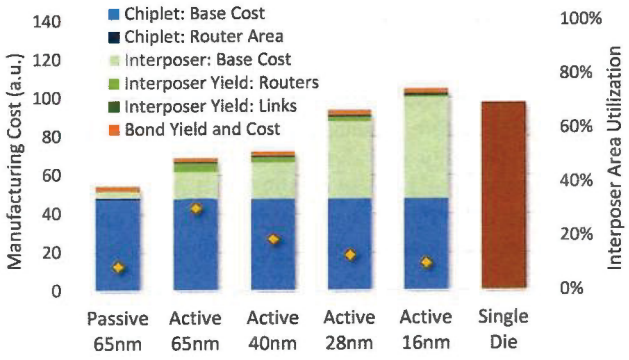


Fig. 12. Manufacturing cost breakdown of 32-core four-chiplet interposers systems versus 16nm monolithic die. Interposer utilization on secondary axis.

latency and high bandwidth, a NoC with link width of 256 bits with one router per core is used for the eight-core circuit, and a link width of 512 bits with eight routers (two per chiplet, with local networks within each four core cluster), both in a 2×4 mesh.

Using the equations in Section II-A and IV-A, with publicly licensable industry wafer and TSV costs from *IC Knowledge, LLC* [7], we calculate the total system costs for a selection of interposer process technologies, with results shown in Figures 11 and 12. The interposer is assumed to add a 10% area overhead for space around the chiplets. For each interposer, the cost overhead includes the base interposer silicon (as if it had ideal yield), the losses from router and interconnect yields, the bonding cost overhead from process complexity, and pessimistic bond yield of 99%. Additionally, the passive interposer includes a cost overhead for the NoC router area that must be added to the chiplets. The total cost of the chiplets is included for each interposer process, and the resulting systems are compared against the cost to manufacture a monolithic chip. The yields for the chiplets and monolithic die use the core-binning methodology from Section II-B, with the yield defined as the percentage of dies that produce *some* level of functional binning. Thus these results represent that average manufacturing cost per functional system, but do not reflect the improvements from Section II-B to core count and speed binning that come from chiplet partitioning and matching.

As visible from the model results, the interposer price is generally dominated not by the yield of the interconnect and routers, but by the base fabrication cost of the silicon. The most recent process nodes in our analysis at 28nm and 16nm demonstrate increasing price per area, and although router area scales and yield improves, the area of the interposer is constrained by the chiplets and does not shrink. With these recent processes, the base interposer cost outweighs the yield improvements from chiplet partitioning, especially for the smaller eight-core system. However, the passive and active interposers at older processes can be cost-effective even for mainstream systems, and they demonstrate significant reductions for the large area 32-core system, even before core and parametric binning improvements are considered. The passive interposer is consistently lower cost than the active interposer, but a fully-active interposer at a mature process still demonstrates cost-effectiveness while supporting improved network performance and design flexibility. System designers can thus select the proper interposer solution for project-specific performance and cost needs.

C. Interposer Yield Increase through Fault Tolerance

Although the active interposer suffers from increased wafer cost and has vulnerable critical logic area, it is possible to reduce the vacant "prepaid" area, and active devices can be used to improve active interposer functionality. To improve the yield of the active interposer, the free area can be used for fault tolerance methods that virtually remove yield losses. With these techniques, active interposers are no longer constrained to "minimally-active" designs, allowing for cost-effective high-performance NoC integration. These techniques are also applicable to passive interposers, but the area overhead is added to the routers in the chiplets, leading to a cost versus yield trade-off.

1) Fault Tolerant Routers: As shown in Section IV-B, the active router area can be the largest yield concern for active interposers with high bandwidth networks. A range of prior literature has proposed methods to add fault tolerance to NoC routers. In particular, the router design from Wang et al. [27] is a promising candidate that employs low-overhead fault tolerance strategies to the routing computation, virtual channel allocation, switch allocation, and crossbar. The fault tolerance mechanisms add only 27% to the router area while allowing for functional behavior until a mean of 21 defects per router, virtually eliminating router failures in the active interposer with no die size increase. The addition of router fault tolerance is especially important for active interposers in aging process nodes like 65nm, in which the router area may consume a significant percentage of the active interposer. For example, the cost overhead of the 32-core active interposer in 65nm is reduced by 23% through router fault tolerance, with costs just 4% higher than if the interposer achieved ideal yield.

2) Redundant Interconnect: Applying fault tolerance to the active interposer routers greatly improves interposer yield, but the interconnect can still be a point of failure that reduces yield and thus increases interposer cost. Because interconnect

yield is relatively high for wide routes, adding only a small number of redundant wires to each link in the NoC is sufficient for achieving close-to-ideal yield. Based on Equation 5 and the 32-core case study with 512-bit link width, interconnect yield is 97% before redundancy. By adding two redundant routes per bus, any one short or two cuts can be avoided, improving interconnect yield to >99.9%. If the routers already include fault detection mechanisms, the redundant interconnect overhead can be included with little additional overhead.

D. Other Benefits of Interposer Integration

The analysis in the prior section is meant to demonstrate the benefits to manufacturing and economics that can be realized through interposer-enabled chiplet integration, with moderate benefits for mainstream processors and significant improvements for large server processors. However, interposer-based integration also enables other significant design opportunities. Benefits include 1) on-die integration of high bandwidth DRAM memory stacks or emerging resistive nonvolatile memories, 2) heterogeneous processes for analog/RF, high speed SerDes, etc., and 3) chiplet-enabled IP reuse. The potential performance, efficiency, and cost benefits of these technologies is beyond the scope of this work, but the results from this analysis suggest that the overheads of interposer integration are cost-effective when combined with die partitioning, thus enabling these advanced design options free of charge with better performance and efficiency than low-bandwidth integration methods like Package on Package (PoP). Alternatively, a high-performance design that requires an interposer for memory integration should also explore the manufacturing benefits of die partitioning using the methodology described here.

V. Conclusions

As demonstrated with the models developed in this work, interposer-enabled die partitioning can provide significant improvements to the manufacturability of previously unconsidered systems as small as mainstream desktop processors. For a 32-core server processor, yield loss was reduced by 0.42× and the number of fully enabled systems was increased by 1.98×-3.94×, depending on process maturity. Contrary to prior assumptions, these yield improvements can be cost-effectively realized, while still providing high-performance communication, through the use of either active or passive interposers, depending on performance and cost requirements. Active interposers can provide several times lower latency and higher throughput links than passive interposers, but the low wafer cost of passive interposers provides a cost advantage, even after including active interposer fault tolerance methods for improved yield. This work aims to provide system designers with the proper guidelines and tools for determining the best interposer solution to meet system requirements, proliferating the usage of the interposer-chiplet design approach to a broader application range.

Acknowledgments

AMD, the AMD Arrow logo, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.
2017 Advanced Micro Devices, Inc. All rights reserved.

References

[1] AMD Epyc. http://www.amd.com/en/products/epyc.
[2] International technology roadmap for semiconductors 2.0, 2015 edition, system integration. Report Ch 1, Semiconductor Industry Association, 2015.
[3] Anandtech. Cpu benchmarks. http://anandtech.com/bench/CPU/1603.
[4] J. A. Carballo et al. ITRS 2.0: Toward a re-framing of the semiconductor technology roadmap. In *IEEE 32nd Intl. Conf. Computer Design*, Oct 2014.
[5] P. Christie and J. P. de Gyvez. Prelayout interconnect yield prediction. *IEEE Trans. Very Large Scale Integr. Syst.*, 11(1), Feb. 2003.
[6] J. A. Cunningham. The use and evaluation of yield models in integrated circuit manufacturing. *IEEE Trans. Semicond. Manuf.*, May 1990.
[7] IC Knowledge LLC. *IC Cost and Price Model, 2016 Revision 05*, 2016.
[8] N. E. Jerger et al. NoC architectures for silicon interposer systems: Why pay for more wires when you can get them (from your interposer) for free? In *47th IEEE/ACM Int. Symp. Microarchitecture*, Dec 2014.
[9] A. Kannan et al. Enabling interposer-based disintegration of multi-core processors. In *48th Annual IEEE/ACM Intl. Symp. Microarchitecture*, Dec 2015.
[10] J. Kim. Active Si interposer for 3D IC integrations. In *Int. 3D Systems Integration Conf.*, Aug 2015.
[11] N. Kim et al. Interposer design optimization for high frequency signal transmission in passive and active interposer using through silicon via. In *IEEE 61st Electronic Components and Tech. Conf.*, May 2011.
[12] M. Lapedus. 10nm versus 7nm. http://semiengineering.com/10nm-versus-7nm/, April 2016.
[13] M. Lapedus. Battling fab cycle times. http://semiengineering.com/battling-fab-cycle-times/, Feb 2017.
[14] J. H. Lau. TSV manufacturing yield and hidden costs for 3D IC integration. In *60th Electronic Components and Tech. Conf.*, June 2010.
[15] C. C. Lee et al. An overview of the development of a GPU with integrated HBM on silicon interposer. In *IEEE 66th Electronic Components and Tech. Conf.*, May 2016.
[16] S. Li et al. McPAT: An integrated power, area, and timing modeling framework for multicore and manycore architectures. In *42nd IEEE/ACM Intl. Symp. Microarchitecture*, Dec 2009.
[17] K. Low. Samsung foundry's business strategy. http://semiengineering.com/samsung-foundrys-business-strategy/, April 2016.
[18] N. Pantano et al. Technology optimization for high bandwidth density applications on 3D interposer. In *6th Electronic Sys.-Integration Tech. Conf.*, Sept 2016.
[19] K. Saban. Xilinx stacked silicon interconnect technology delivers breakthrough FPGA capacity, bandwidth and power efficiency. https://www.xilinx.com/support/documentation/white_papers/wp380_Stacked_Silicon_Interconnect_Technology.pdf, Dec 2012.
[20] M. J. Schulte et al. Achieving exascale capabilities through heterogeneous computing. *IEEE Micro*, 35(4), July 2015.
[21] C. H. Stapper. Defect density distribution for LSI yield calculations. *IEEE Trans. Electron Devices*, 20(7), Jul 1973.
[22] D. Stow et al. Cost analysis and cost-driven IP reuse methodology for SoC design based on 2.5D/3D integration. In *IEEE/ACM Intl. Conf. Computer-Aided Design*, Nov 2016.
[23] D. Stow et al. Cost and thermal analysis of high-performance 2.5D and 3D integrated circuit design space. In *IEEE Computer Society Annu. Symp. VLSI*, July 2016.
[24] S. Sutardja. 1.2 the future of IC design innovation. In *IEEE Int. Solid-State Circuits Conf.*, Feb 2015.
[25] D. Velenis et al. Processing active devices on Si interposer and impact on cost. In *Int. 3D Systems Integration Conf.*, Aug 2015.
[26] P. Vivet et al. 3D advanced integration technology for heterogeneous systems. In *Int. 3D Systems Integration Conf.*, Aug 2015.
[27] L. Wang et al. A high performance reliable NoC router. In *21st Asia and South Pacific Design Automation Conf.*, Jan 2016.
[28] G. Yeric. Moore's law at 50: Are we planning for retirement? In *IEEE Intl. Electron Devices Meeting.*, Dec 2015.
[29] W. Zhao and Y. Cao. Predictive technology model for nano-CMOS design exploration. *ACM J. Emerging Tech. Computing Systems*, Apr. 2007.