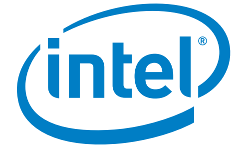
WHITE PAPER

FPGA



Enabling Next-Generation Platforms Using Intel's 3D System-in-Package Technology

Next-generation platforms increasingly require higher bandwidth, flexibility, and functionality while lowering power profiles and footprint requirements.

Author

Manish Deo
Senior Product Marketing Manager
Intel Programmable Solutions Group

Table of Contents

Introduction 1

Next-generation system challenges 1

In-package integration using Intel's heterogeneous 3D SiP technology 3

Conclusion 6

References 7

Where to get more information . . . 7

Introduction

This white paper evaluates system requirements for next-generation platforms and explains why conventional solutions may no longer be able to meet these requirements effectively. The paper introduces the heterogeneous 3D system-in-package (SiP) technology featured in Intel® Stratix® 10 FPGAs and SoCs. This technology enables next-generation platforms by offering higher bandwidth, lower power, a smaller form factor, and increased functionality and flexibility. Stratix 10 FPGAs and SoCs feature 3D SiP-based transceivers across all densities. This paper elaborates on the scalability, flexibility, and fast time-to-market benefits of this next-generation transceiver solution. Additionally, it enumerates the physical construct of the SiP technology, compares it with alternative approaches, and explains how this technology addresses specific requirements for next-generation platforms.

Next-generation system challenges

Next-generation platforms are evolving rapidly to keep pace with emerging system trends driven by an explosion of applications such as data center capabilities, Internet of Things (IoT), 400G to terabit networking, optical transport, 5G wireless, 8K video, etc. The resulting expansion of connectivity and processing will affect the semiconductor space significantly, from the type of components that are built to higher efficiency systems and related services. A close evaluation of this emerging landscape reveals some interesting trends.

For example, next-generation data center workloads demand increasingly higher computational capabilities, flexibility, and power efficiencies; outstripping the capabilities of today's general-purpose servers. Additionally, data center infrastructure must be virtualized and delivered as a service over commodity servers to reduce complexity and provide greater business agility and scalability. However, server performance improvements have actually slowed, primarily due to power limitations. Designing data center solutions for specific workloads increases efficiency but significantly limits the homogeneity and flexibility of the solution. Flexibility is crucial because data center services evolve rapidly and require adaptable hardware. As a result, the challenge for next-generation data center platforms is to deliver higher performance (acceleration), power efficiency, and flexibility simultaneously.

IoT reflects similar challenges. IoT is projected to grow dramatically and hit the multi-billion "smart objects" mark in the near future. These smart objects are connected and communicate with each other or to a cloud or data center. The infrastructure must determine which data needs to be processed and which data is dropped, all in real time. Therefore, IoT requires a highly connected, flexible, efficient, bandwidth-rich infrastructure that enables insight from the data center to

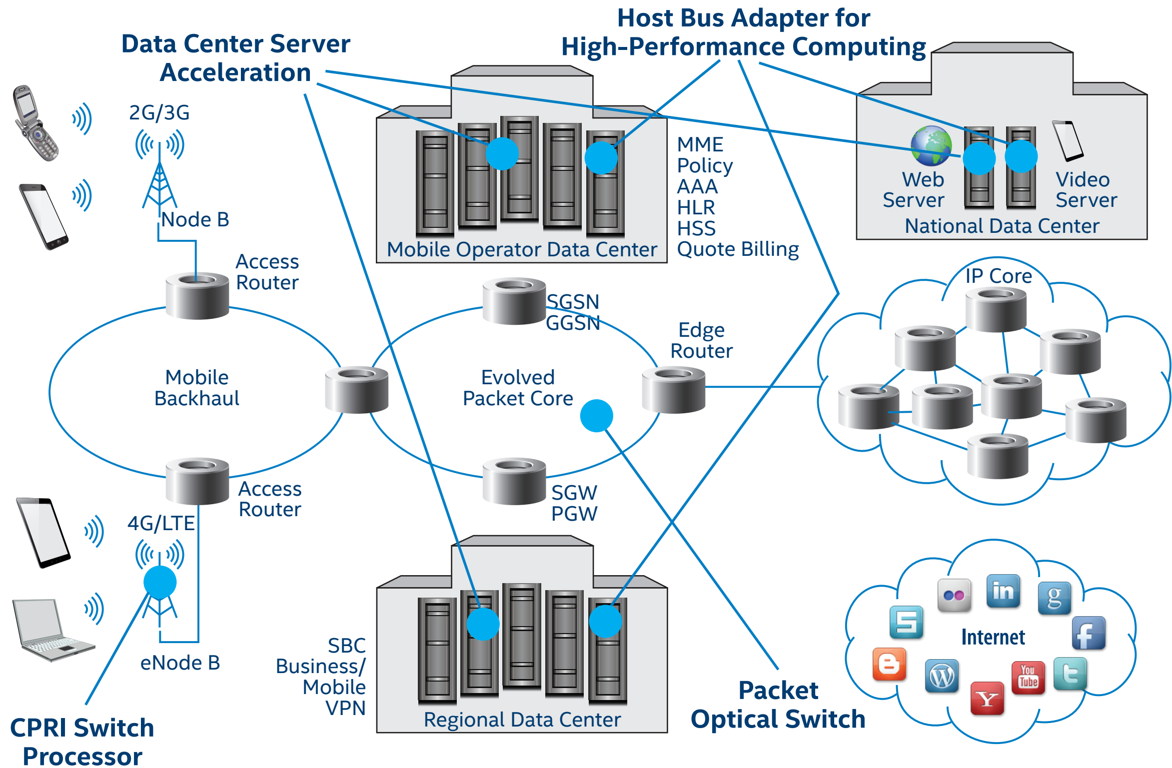


Figure 1. Data Center Acceleration

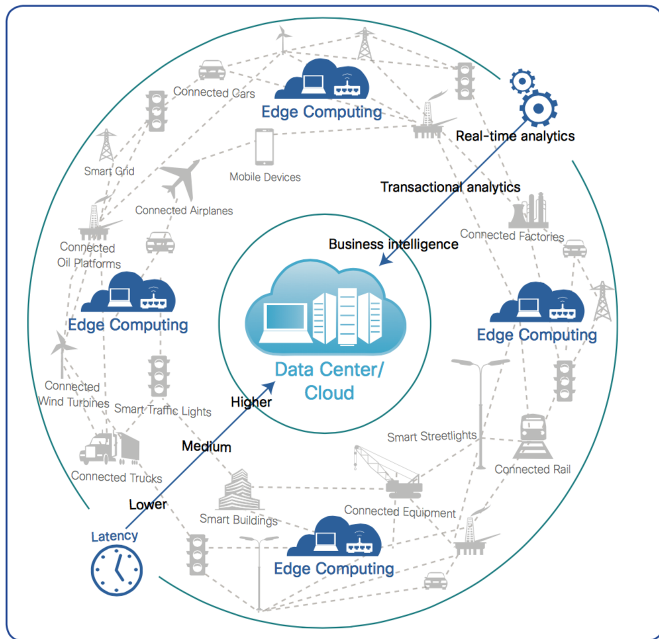


Figure 2. IoT Landscape

the edge. This requirement challenges service providers, data centers, cloud computing, and storage systems to satisfy this insatiable demand for Internet traffic.

Next-generation platforms reflect a common underlying theme: the need for increased bandwidth and functionality with lower power profiles and footprint requirements. Simply put, the devices used to build these next-generation platforms must do more, be faster, take less printed circuit board (PCB) real estate, and burn less energy, all at the same time. This challenge requires innovative solutions across the semiconductor ecosystem.

Thus, system architects designing next-generation platforms must try to meet the following requirements:

- Higher bandwidth
- Lower power
- Smaller footprint or form factor
- Increased functionality
- Increased flexibility

Historically, system architects have responded to these requirements by packing more discrete components on a standard PCB, attempting to provide maximum functionality and performance while keeping power budgets in check. Figure 3 depicts a conventional solution in which multiple discrete components (FPGA, memory, CPU, etc.) are assembled on a standard PCB.

This conventional integration scheme is nearing its logical end as it struggles to keep pace with-next generation requirements. Some of the key challenges are:

- Chip-to-chip bandwidth limited by the interconnect density permitted by underlying PCB.
- System power is too high due to the need to drive long PCB traces between components.
- Form factor is too big due to the number of discrete components required for the desired system functionality.

System architects have looked at monolithic integration for some components to address these limitations. However, this integration leads directly to another challenge: IP maturity. Figure 4 illustrates this conundrum. Different IP blocks

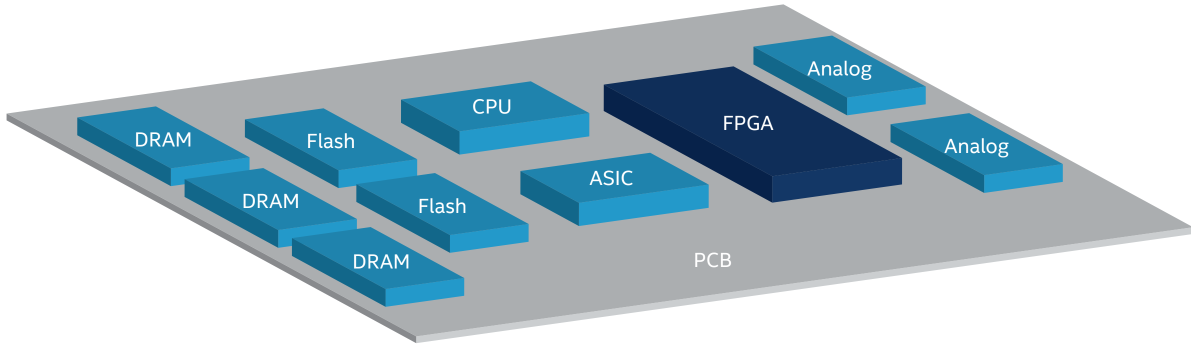


Figure 3. Conventional Discrete Component Integration using PCB

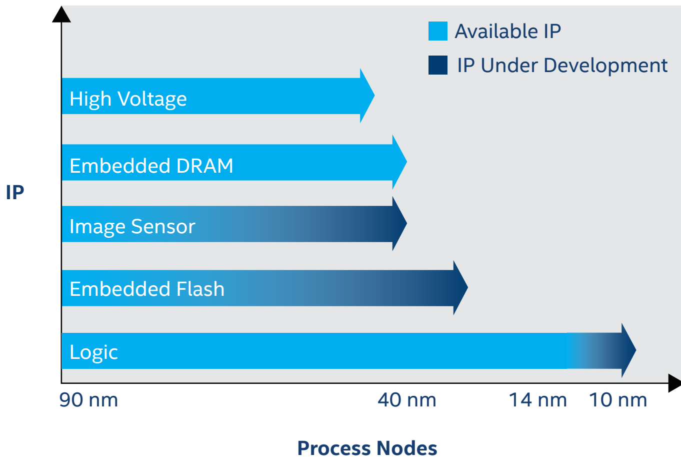


Figure 4. IP Maturity vs. Process Nodes Example

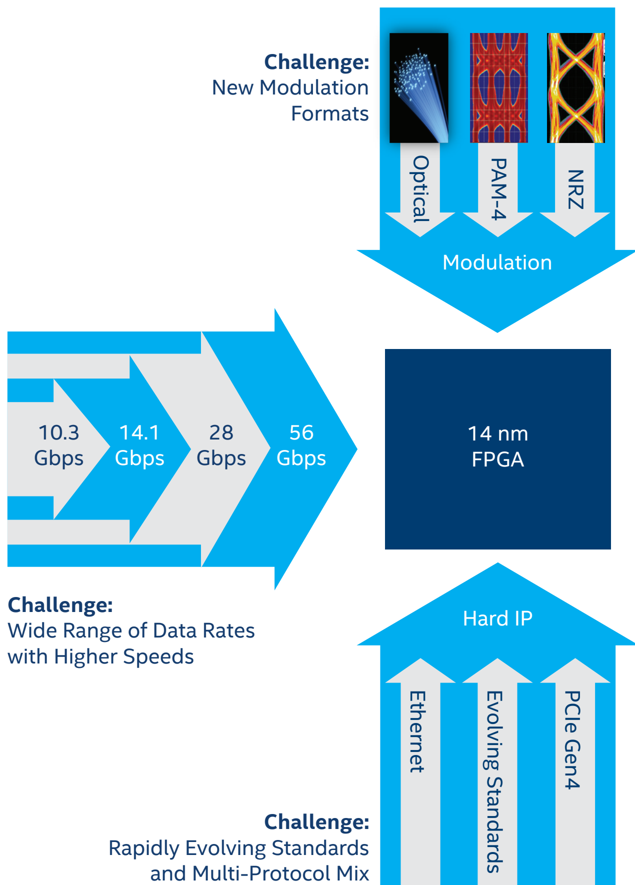


Figure 5. IP Emerging Challenges when Providing Maximum High-Speed Connectivity

mature at different process nodes, and by extension are available at different times. Therefore, it is not possible to integrate all desired IP blocks or functionality monolithically. For example, if a vendor is building a logic die using 14 nm technology and wants to integrate DRAM on chip, the only option is DRAM built using 40 nm or older technology. This limitation does not facilitate a monolithic solution.

Another key challenge is the need to provide maximum high-speed connectivity between devices. FPGA vendors have historically responded by leveraging cutting-edge transceiver technology. Intel industry-leading transceiver technology currently supports data rates over 28 Gbps. Next-generation devices such as Stratix 10 FPGAs and SoCs build on this leadership and plan to support data rates up to 56 Gbps. As shown in Figure 5, modulation schemes for high-end data rates are evolving and newer schemes such as PAM-4 are increasingly promising. Additionally, devices require more embedded hard protocol IP to meet customer requirements. However, many of these protocol standards continue to evolve. Thus, it is imperative to define an innovative solution that integrates emerging technologies and IP blocks quickly.

The challenges posed by next-generation systems have begun to define the solution landscape. Conventional solutions cannot meet the requirements of the future: higher bandwidth, lower power, smaller form factor, and increased functionality and flexibility. The challenge is to develop an innovative, commercially viable, scalable solution that meets these requirements.

In-package integration using Intel's heterogeneous 3D SiP technology

With Stratix 10 FPGAs and SoCs, Intel introduces the heterogeneous 3D system-in-package (SiP) technology. This unique solution addresses all of these challenges: higher bandwidth, lower power, smaller form factor, and increased functionality and flexibility. It also enables in-package integration that is scalable and straightforward to manufacture. This solution combines the right mix of functionality on the right process nodes to provide the system functionality customers need in a single package. The heterogeneous 3D SiP technology enables in-package integration of a range of components such as analog, memory, ASIC, CPU, etc. (see Figure 6). It also integrates transceiver die or tiles from different process nodes in the same package. The following sections describe how Intel's heterogeneous 3D SiP technology combines a monolithic fabric with transceiver tiles.

3D SiP-based transceiver tiles: maximizing scalability and flexibility

Heterogeneous 3D SiP technology decouples the transceiver tile or die from the core fabric die: the transceiver is separate and sits next to the core fabric die. Therefore, the transceiver and core fabric die do not have to be manufactured on the same process node. Figure 7 shows this integration conceptually.

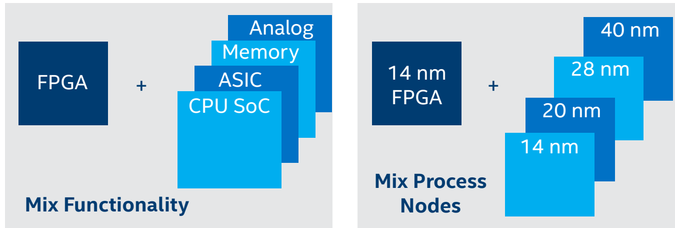


Figure 6. Heterogeneous In Package Integration with 3D SiP Technology

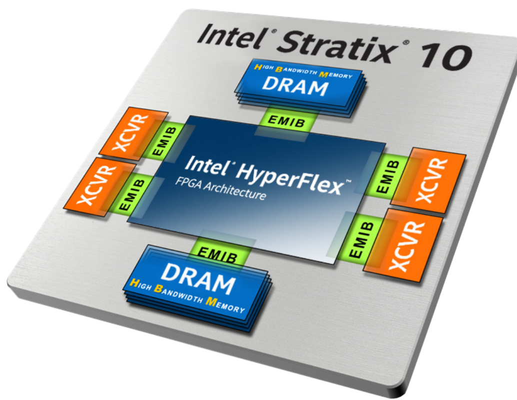


Figure 7. Conceptual Representation of Intel's Heterogeneous 3D SiP Technology

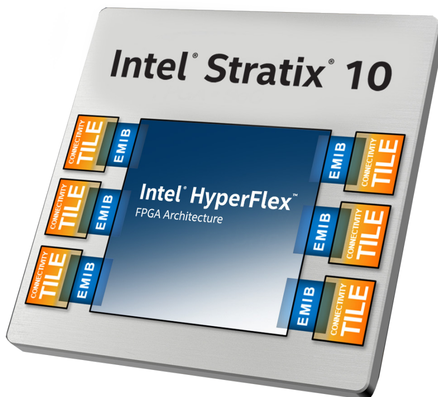


Figure 8. Enhanced Flexibility and Scalability with Separate Transceiver Tiles

The heterogeneous 3D SiP technology allows Intel to mix components to match system requirements, effectively providing robust solutions more quickly than in previous generations. Stratix 10 devices will leverage proven transceiver IP, significantly reducing validation and bring up times, and dramatically improving customers' time-to-market metrics. Looking forward, the 3D SiP technology provides a scalable solution that is capable of supporting 56 Gbps transceivers with emerging modulation schemes such as PAM-4. Similarly, separate transceiver tiles enable support for custom embedded IP. For example, initial Stratix 10 transceiver tiles include a PCIe* Gen3 x16 hard IP block. Future versions could potentially support a variety of hard IP modules, such as PCIe Gen4, multi-port Ethernet, optical, etc., as shown in Figure 8.

The technology that makes broad deployment possible is Intel's patented, state-of-the-art Embedded Multi-die Interconnect Bridge (EMIB). Intel designed EMIB for solutions that require advanced packaging and test capabilities. The EMIB technology provides a simple integration flow and offers an ultra-high density interconnect between heterogeneous die in the same package. It also enables in-package functionality that was either too complex or too cost prohibitive to implement with alternative in-package integration solutions. As illustrated in Figure 7, the EMIB connects the transceiver die to the monolithic FPGA fabric. Next, this paper takes a deeper look at the EMIB technology and the significant advantages it offers compared to alternative integration solutions.

The EMIB advantage

The EMIB technology offers a simpler manufacturing flow, higher performance, enhanced signal integrity, and reduced complexity. Figure 9 depicts the physical package construction. The construction heterogeneously integrates the FPGA fabric die (1) and two transceiver die (2). (The non-FPGA die can be a transceiver die, memory die, CPU die, or any other functionality.) The three die rest on a standard flip-chip ball grid array (FCBGA) package substrate (3), which connects to the underlying PCB. The routing between the die and the package balls uses standard FCBGA routing (4). This assembly is encapsulated with a standard package lid (5) to create a single package solution. The package substrate utilizes several EMIB connections (6). The EMIB enables the heterogeneous in-package integration by connecting the die using an ultra-high-density interconnect.

Innovative, smarter in-package integration

As illustrated in Figure 10, the EMIB is a small silicon chip embedded in the underlying package substrate and offers dedicated ultra-high-density interconnect between die. Importantly, the EMIB physical dimensions do not limit the number of die that can be integrated. In contrast, alternative implementations use a large piece of silicon interposer that sits on top of the package substrate and exceeds the entire length of the die to be integrated. The large piece of silicon interposer makes the solution cost prohibitive and prone to issues such as warpage, etc. Alternative solutions also

require a large number of micro bumps using micro vias, which affects the overall yield and manufacturing complexity. Additionally, the number of die that can be integrated using an interposer is limited, affecting the scalability.

Higher performance

Stratix 10 FPGAs and SoCs leverage the EMIB's heterogeneous in package integration capabilities to offer the highest levels of performance. As shown in Figure 10, the EMIB enables the die I/O or bumps to be integrated to be placed as close as possible to the edge of the die because fewer I/O or bumps are required. This methodology ensures that the physical connections between the die are very precise and use short interconnect wires. The short wires, in turn, result in significantly reduced loading that the wire presents to the driving buffer, resulting in higher performance. In contrast, alternative solutions reconnect the logic fabric using the large underlying interposer. This homogeneous integration involves a connecting large number of I/O or bumps, which spreads them out from the edge of the die towards the center. This placement results in much longer interconnect wires and higher loading on the driving buffers. The net result is lower performance.

Reduced complexity, superior signal and power integrity

The EMIB-based flow significantly reduces manufacturing complexity. As illustrated in Figure 11, the EMIB solution offers simple two-step connectivity for user I/O, power, and transceiver signals: bump to standard package trace to package ball. The standard package trace is widely used in FCBGA packages. This simple connectivity results in reduced manufacturing complexity and superior signal and power integrity. Key parameters such as insertion loss to crosstalk ratio (ICR) and power supply rejection ratio (PSRR) are comparable to monolithic designs.

In contrast, alternative solutions offer complex four-step connectivity for signals that need to connect to the package ball: bump to through silicon vias (TSV) to bump to package trace to package ball. This connectivity requires TSVs for every user signal and adds significant complexity to the manufacturing flow. (The EMIB flow does not use any TSVs.) TSV processing adds significant manufacturing complexities leading to incremental yield loss, which affects the overall commercial viability. Additionally, alternative solutions use large number of TSVs (~10,000). This complex four-step connectivity results in poor signal integrity for high-speed signals and causes IR drop for power delivery nets. TSVs also add series resistance and capacitance, which makes high-speed design for the transceiver blocks even more complex and challenging. Cross talk in the interposer routing and coupling between TSVs may impact ICR specifications; signal-to-power rail coupling through TSVs may impact PSRR specifications.

While heterogeneous 3D SiP integration provides an ideal solution to address scalability and flexibility requirements for next-generation transceivers, peripherals, memories etc., a monolithic FPGA fabric is vital in meeting next-generation platform requirements. The next section will detail the benefits of using a monolithic fabric versus interposer based stacked core fabric solution.

Monolithic core fabric: maximum performance and utilization

A monolithic FPGA core fabric is critical to provide maximum performance and utilization, and ensure that data can be processed at the highest rates possible without running into

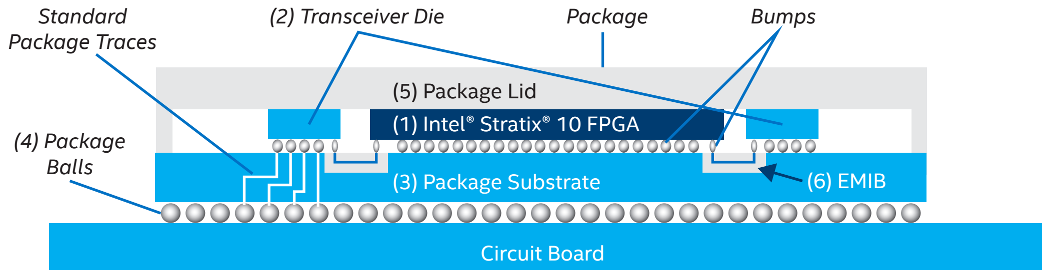
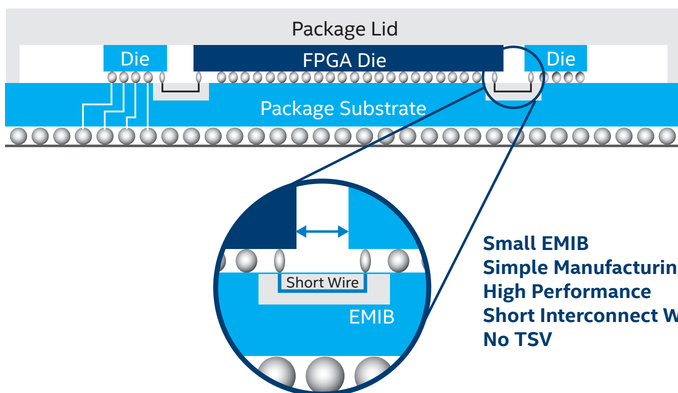


Figure 9. Heterogeneous Integration using EMIB Technology



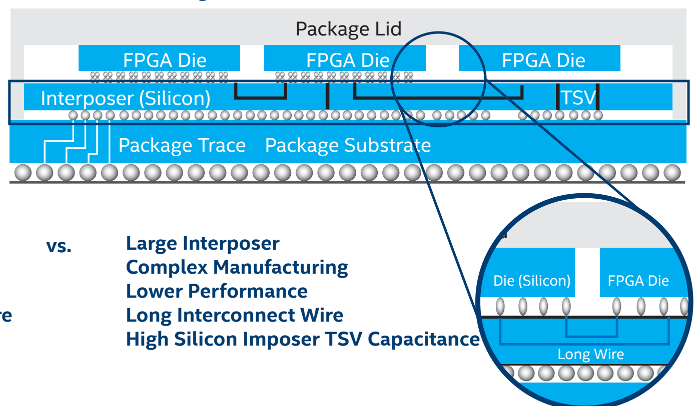
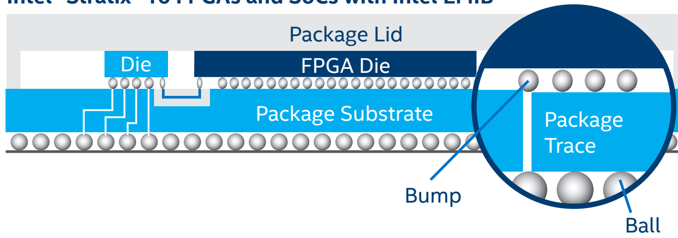


Figure 10. EMIB Implementation vs. Alternative Interposer based Implementation



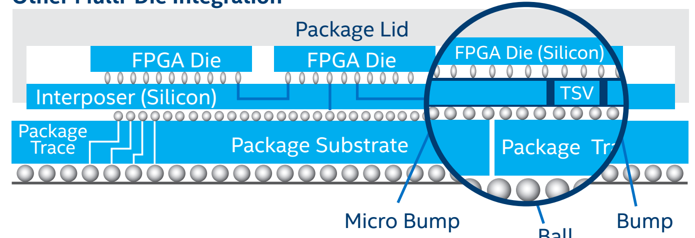


Figure 11. Reduced Complexity and Superior Signal Integrity with EMIB

routing congestion, utilization bottlenecks, or degraded performance. With 14 nm Tri -Gate process technology and the new Intel HyperFlex™ FPGA Architecture, the FPGA fabric also delivers 2X performance gains on average compared to previous generations. † Figure 12 compares a device with monolithic fabric such as Stratix 10 FPGAs with competing products that rip apart the monolithic fabric in multiple slices and then reconnect them using interposer-based techniques.

Alternative solutions suffer from severe performance degradation compared to a monolithic fabric. Figure 13 shows the results of independent research by researchers at the University of Toronto and University of Sao Paulo.

- The diamond, circle, and triangle lines represent 0.5, 1.0, and 1.5 ns interposer delays experienced by alternative solutions that attempt to connect logic using conventional interposers.
- The triangle line represents interposer delays incurred by alternative solutions available on the market today.
- The square line represents a monolithic fabric, which has no interposer delay.

The results show that there is a significant performance degradation incurred when connecting logic fabric using interposers. In fact, the results show up to 50% performance degradation for a 1.5 ns interposer delay relative to a purely monolithic solution.

- For more information on this research, refer to the paper CAD and Routing Architecture for Interposer-Based Multi-FPGA Systems by Andre Hahn Pereira, University of Sao Paulo and Vaughn Betz, University of Toronto.

Conclusion

Next-generation platforms increasingly require innovative solutions that provide significantly higher performance, lower power, and smaller form factors. The explosion in data center capabilities and proliferation of IoT technologies are emerging as key drivers. In addition, advancements in terabit networking, optical transport, 8K video, and 5G wireless domains are ramping up rapidly, forcing the semiconductor ecosystem to find innovative solutions.

Stratix 10 FPGAs and SoCs enable a wide range of next-generation platforms by facilitating products that provide significantly higher bandwidth, lower power, smaller form factor, and increased functionality and flexibility. Stratix 10 FPGAs and SoCs leverage Intel's patented EMIB technology to enable effective in-package integration of system-critical components such as analog, memory, ASICs, CPUs, etc. This solution offers increased scalability, lower risk, reduced time to market, and addresses a wide range of applications. Additionally, Stratix 10 FPGAs and SoCs combine Intel's

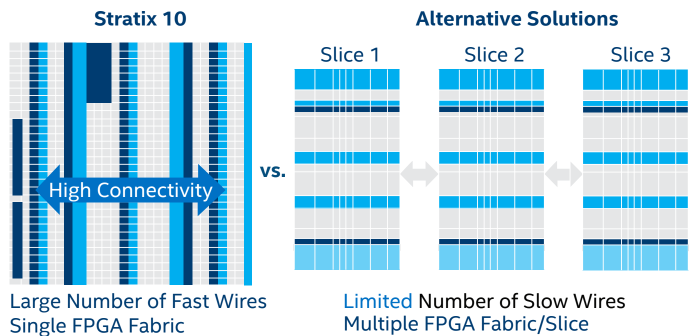


Figure 12. Stratix 10 Device with Monolithic Fabric vs. Alternative Stacked Core Fabric Solution

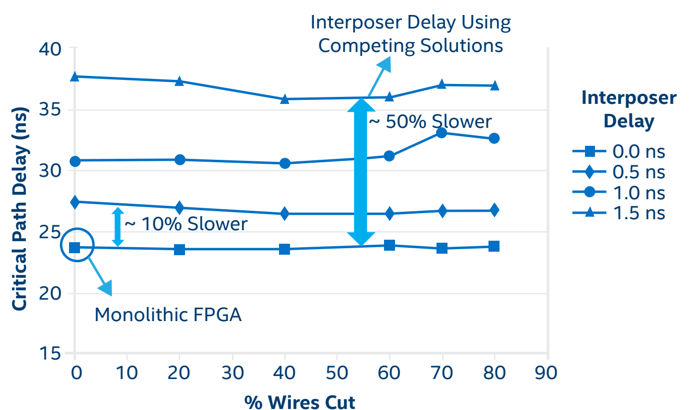


Figure 13. Performance Degradation Relative to Monolithic Fabric

14 nm Tri-Gate process technology and the new Intel HyperFlex FPGA Architecture to provide 2X performance gains, on average, relative to previous generations. † The confluence of process technology (14 nm Tri-Gate), monolithic core fabric (with advanced Intel HyperFlex FPGA Architecture), and state-of-the-art package integration in Stratix 10 FPGAs and SoCs promises to revolutionize the landscape for next-generation platforms.

References

[1] http://www.intel.com/content/www/us/en/foundry/emib.html

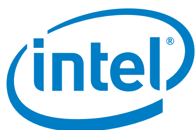
[2] Microsoft White Paper: A Reconfigurable Fabric for Accelerating Large-Scale Datacenter Services, Andrew Putnam, Adrian M. Caulfield, Eric S. Chung, Derek Chiou, Kypros Constantinides, John Demme, Hadi Esmaeilzadeh, Jeremy Fowers, Gopi Prashanth Gopal, Jan Gray, Michael Haselman, Scott Hauck, Stephen Heil, Amir Hormati, Joo-Young Kim, Sitaram Lanka, James Larus, Eric Peterson, Simon Pope, Aaron Smith, Jason Thong, Phillip Yi Xiao, Doug Burger.

[3] Cisco White Paper: Attaining IoT Value: How To Move from Connecting Things to Capturing Insights, Gain an Edge by Taking Analytics to the Edge, Andy Noronha, Robert Moriarty, Kathy O'Connell, Nicola Villa.

Where to get more information

For more information about Intel and Stratix 10 FPGAs, visit **https://www.altera.com/products/fpga/stratix-series/stratix-10/overview.html**

† Tests measure performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.



Intel, the Intel logo, the Intel Inside mark and logo, Altera, HyperFlex, and Stratix are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.
* Other marks and brands may be claimed as the property of others.
© Intel Corporation.

Please Recycle