WHITE PAPER

FPGA

Intel® Stratix® 10 MX Devices with Samsung* HBM2 Solve the Memory Bandwidth Challenge

The Intel Stratix 10 MX device family helps customers efficiently meet their most demanding memory bandwidth requirements.

SAMSUNG

Authors

Manish Deo
Senior Product Marketing Manager
Intel Programmable Solutions Group

Jeffrey Schulz
In-Package I/O Implementation Lead
Intel Programmable Solutions Group

Lance Brown
Senior Strategic and
Technical Marketing Manager
Intel Programmable Solutions Group

Table of Contents

Introduction

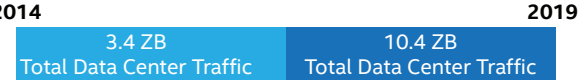
Conventional memory solutions have limitations that make it difficult for them to meet next-generation memory bandwidth requirements. This white paper describes the emerging memory landscape that addresses these limitations. The Intel® Stratix® 10 MX DRAM system-in-package (SiP) family combines a 1 GHz high-performance monolithic FPGA fabric, state-of-the-art Intel Embedded Multi-die Interconnect Bridge (EMIB) technology, and High Bandwidth Memory , all in a single package. The Intel Stratix 10 MX family helps customers efficiently meet their most demanding memory bandwidth requirements, which are not possible using conventional memory solutions. This white paper describes the solution and explains how Intel Stratix 10 MX devices solve the memory bandwidth challenge for key end markets and applications.

The Memory Bandwidth Challenge

Memory bandwidth is a critical bottleneck for next-generation platforms. The critical path in any system's performance is the system's ability to process large amounts of data quickly. Computing elements (such as the FPGA or CPU) must read or write large chunks of data to and from memory efficiently.

A host of end markets and applications (including data center, high-performance computing systems, broadcast (8K), wireline networking, data analytics, and Internet of Things (IoT)) drive memory bandwidth requirements, and process an ever-increasing amount of data. Data centers and networking platforms channel this data, and strive to increase efficiency and accelerate workloads simultaneously to keep up with the bandwidth explosion. Total data center traffic is projected to hit 10.4 zetabytes (ZB) in 2019.

To meet exploding memory bandwidth requirements, system designers have attempted to use currently available conventional technologies. However, these conventional technologies pose a number of challenges.



1 zettabyte (ZB) is equal to sextillion bytes or a trillion gigabytes

Source: Cisco VNI Global IP Traffic Forecast, 2014 - 2019

Figure 1. Data Center Traffic Projection by 2019

Challenge #1: Limited I/O Bandwidth

I/O bandwidth scaling has not kept up with bandwidth requirements. Simply put, it is not physically possible to have enough I/O pins to support a wide enough memory bus delivering the required bandwidth. Adding more components cannot solve the problem due to the increased power and form factor impact.

The wireline networking market provides a good example of this particular challenge. The aggregate full duplex bandwidth required for packet storage and retrieval is a function of the traffic load. System designers must adequately provision enough margin to ensure sustained line rate performance (100G, 200G, etc.). Figure 2 illustrates the 200G traffic load inflection point in the wireline networking space. An application requires over 700 I/O pins and five DDR4 (x72, 3200 Mbps) DIMMs to meet basic data plane memory functionality. As Figure 2 shows, 400G systems will need over 1,100 I/O pins and 8 DDR4 (x72, 3200 Mbps) DIMMs. Going forward, it is not feasible to increase package I/O counts to meet these bandwidth requirements.

Challenge #2: Flat Power Budgets

Power budgets are a key challenge area. To meet memory bandwidth requirements, system designers must assemble an increasing number of discrete memory (components, DIMMs, etc.) that connect to computing elements using standard PCB traces. A standard DDR4 x72 bit interface consumes roughly 130 parallel PCB traces. Large I/O buffers drive these long PCB traces, thereby consuming significant energy/bit. Additionally, an application requiring 256 GBps memory bandwidth, needs an estimated 10 DDR4 3,200 Mbps DIMMs consuming an estimated 40 W total power. [1] The solution quickly becomes power limited for applications requiring the highest bandwidth levels. On the other hand, system-level power budgets have stayed flat or even dropped. System designers are increasingly challenged to obtain the highest bandwidth per watt at a system level.

[1] Assuming 30% RD/WR, single rank configuration. I/O and controller power included.

Challenge #3: Smaller Form Factor

Attempting to meet increasing memory bandwidth requirements using conventional technologies often means assembling more discrete memory devices on the PCB. To ensure proper system level margins, designers use specific board layout guidelines for trace lengths, termination resistors, and routing layers. These rules limit how compact the design can be, and how closely the devices can be placed. As memory bandwidth requirements grow, it will be increasingly difficult to meet memory bandwidth targets while satisfying form factor constraints using conventional solutions such as DDR. Figure 3 illustrates the form factor impact as the memory bandwidth requirement goes from 80 GBps to 256 GBps.

Challenge #4: JEDEC DDR Bandwidth Sustainability

Conventional technologies such as DDR are finding it increasingly difficult to scale to meet future memory bandwidth requirements. DDR technologies have scaled every generation for the past decade and this scaling is

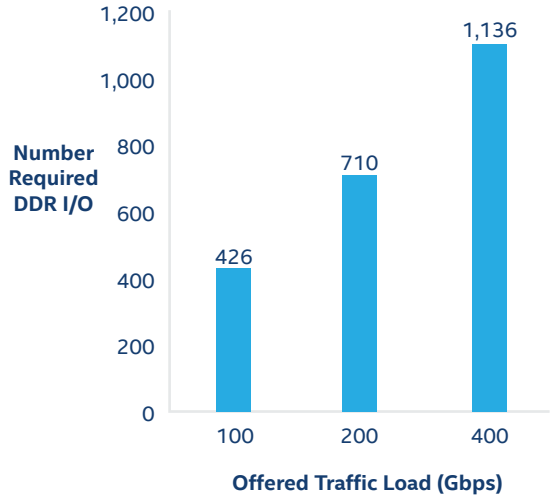
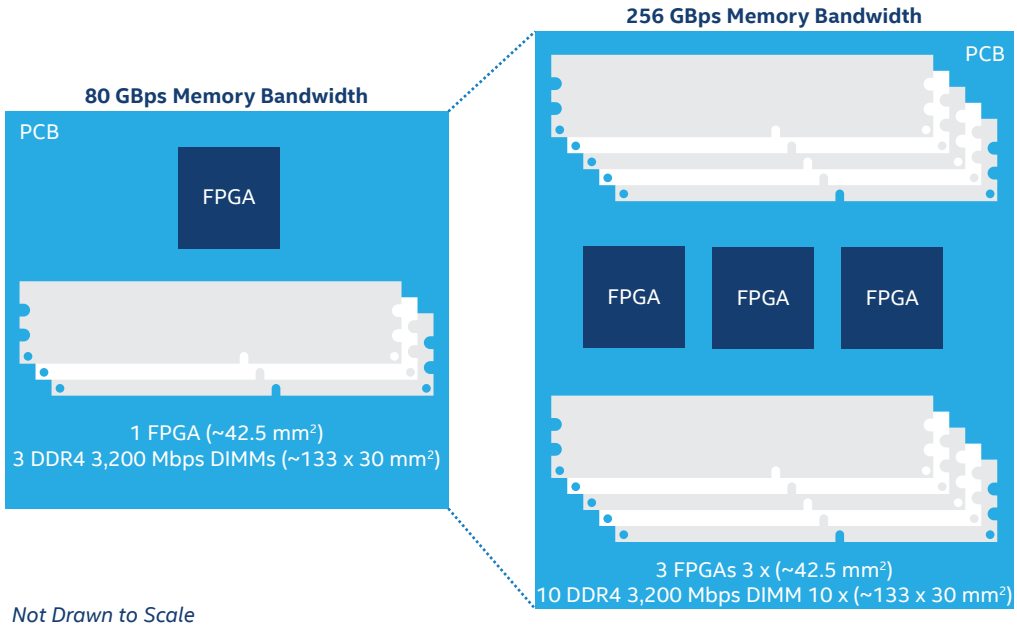


Figure 2. I/O Bandwidth Limited



Not Drawn to Scale

Figure 3. Form Factor Limitations with Conventional Solutions

nearing its end. Projecting out beyond DDR4 and assuming a continued 2X rate of increase, one can estimate bandwidth (per DIMM) in the range of 40 GBps (DDR5). More importantly, the memory bandwidth requirements for next-generation applications are expected to far exceed the trends from the past decade. See Figure 4.

DRAM Memory Landscape

The memory industry has recognized that conventional technologies cannot meet future bandwidth requirements. As a result, the memory landscape is evolving, with multiple competing solutions attempting to address the challenge.

As shown in Figure 5, basic requirements span both control and data plane memories. Control plane or fast path memories (such as SRAM) typically provide the highest performance (random transaction rates) and low latency. Data plane memories (such as DRAM) have high capacity and bandwidth.

Legacy products include standard DDR-based memory solutions. 3D-based memory solutions evolved to meet the high bandwidth, low power, and small form factor challenges. These solutions stack multiple DRAM layers using through-silicon via (TSV) technology. Because the memories are stacked vertically, 3D memory solutions provide maximum capacity with a smaller form factor.

The following 3D memories communicate with the computing element using high-speed serial transceivers or high-density parallel GPIO:

- Hybrid Memory Cube (HMC) is a 3D DRAM memories with a serial interface.
- MoSys bandwidth engine (BE) is a DRAM memory with a serial interface.
- High Bandwidth Memory is a 3D DRAM memory with a parallel I/O interface.

Figure 6 plots power efficiency as a function of bandwidth for various memory types.

- **Low bandwidth/highest power efficiency**—Low-power DDR (LPDDR) provides the highest power efficiency and is a great fit for mobile end markets.
- **Medium bandwidth/medium power efficiency**—DDR3 and DDR4 have been the workhorses of the memory landscape for more than a decade. Wide I/O 2 (WIO2) uses 3D stacking to stack the memory on top of the computing element, providing higher bandwidth with superior power efficiency.

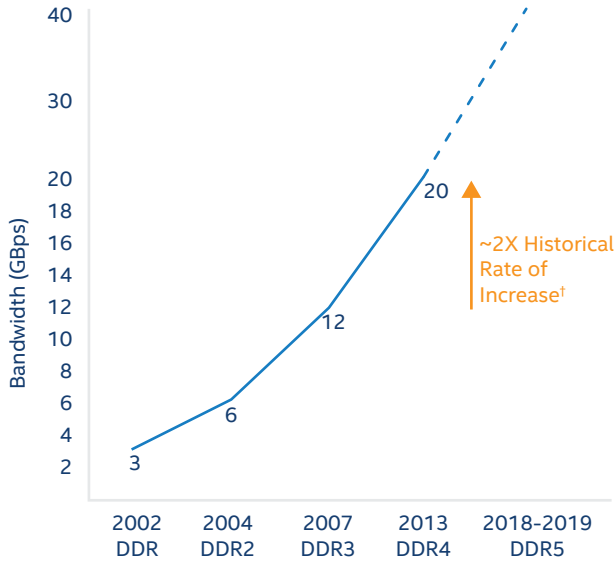


Figure 4. DDR (DIMM) Bandwidth Projection

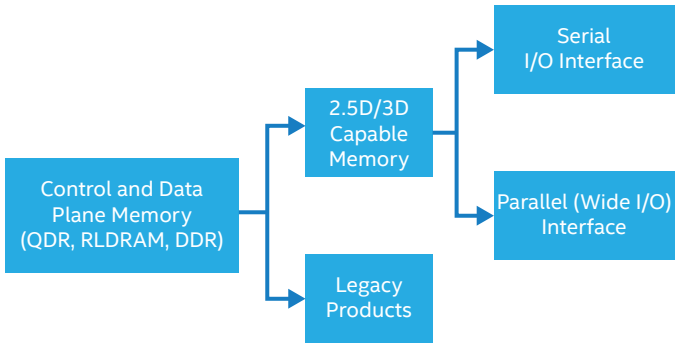


Figure 5. Emerging Memory Landscape

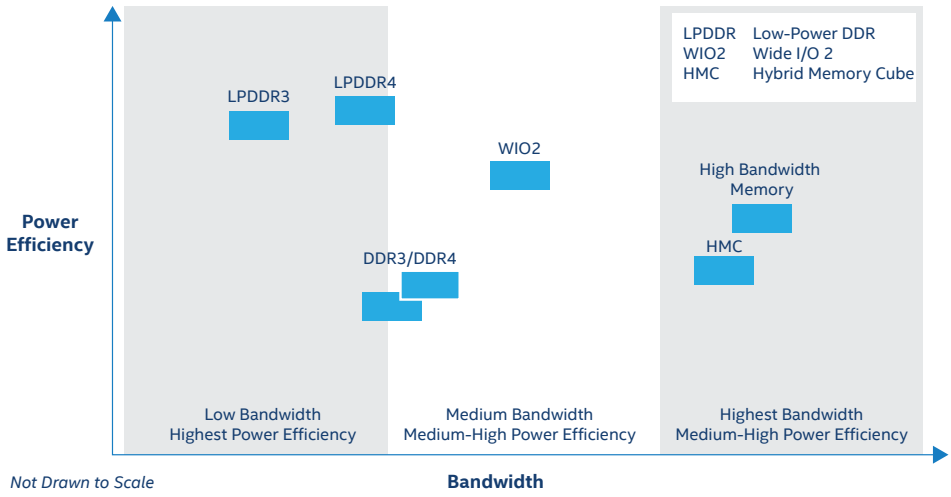


Figure 6. Comparing Power Efficiency vs. Bandwidth

- **High bandwidth/medium to high power efficiency**—High Bandwdith Memory (such as Samsung HBM2) and HMC are competing new technologies.

Intel Stratix 10 MX (DRAM SiP) Devices

A new class of product, Intel Stratix 10 MX devices integrate Samsung* HBM2 with high-performance FPGAs. Intel Stratix 10 MX devices were designed to meet high-performance system demand (where bandwidth is paramount), providing 10X higher memory bandwidth and the highest performance per watt compared to conventional solutions such as DDR.† Figure 7 shows the basic construction of this new product.

Figure 7 (A), the DRAM boxes, show the Samsung HBM2 memory stacks that the device integrates in package (up to 4 stacks integrated per package). Each Samsung HBM2 memory stack is either 4 or 8 memory layers high and supports up to 16 independent channels (64 bits each). Each channel can run at data rates of up to 2 GBps and provide up to 16 GBps of aggregate bandwidth per channel. Figure 8 shows a logical representation of the memory channels and base die.

Figure 7 (B), represent high-performance transceiver tiles that connect to the monolithic core fabric using Intel's EMIB technology.

Figure 7 (C) represents a high-performance monolithic core fabric built using the Intel Hyperflex™ FPGA Architecture. This core fabric can run up to 1 GHz and provide up to 2X performance gains compared to previous generation high-end FPGAs. The high-performance monolithic core fabric ensures efficient processing of the in-package memory bandwidth and enables a viable system-level solution.

Intel EMIB technology enables enables effective in-package integration of different tiles alongside a high-performance monolithic core fabric. The EMIB interface supports the required interface data rates between the core fabric and the Samsung HBM2 memory stack tile. This interface is compatible with standard JEDEC and IEEE 1500 specifications. Intel's EMIB provides an elegant way of integrating multiple tiles in a single package.

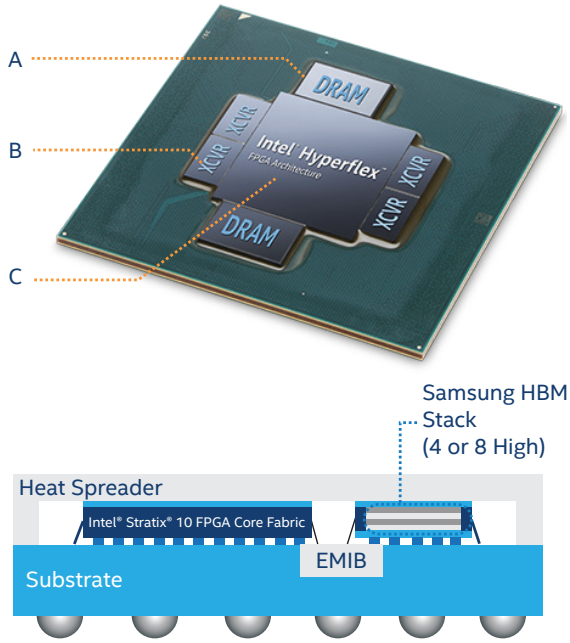


Figure 7. Intel Stratix 10 MX Device

Higher Memory Bandwidth

Integrating up to four Samsung HBM2 stacks and a high-performance FPGA fabric in a single package, Intel Stratix 10 MX devices effectively address the memory bandwidth challenge. Each Samsung HBM2 stack provides up to 256 GBps of aggregate bandwidth. Intel Stratix 10 MX devices thus provide up to 512 GBps aggregate bandwidth in a single package. This unprecedented bandwidth enables multiple applications such as machine learning, data analytics, image recognition, workload acceleration, 8K video processing, and high-performance computing. Intel Stratix 10 MX devices enable solutions that are not possible using conventional memory solutions such as DDR.

Figure 9 compares implementations for an application targeting 400 GBps memory bandwidth. As illustrated, conventional memory solutions using DDR technology will simply not be a feasible option. Intel Stratix 10 MX devices provide an effective implementation and the highest bandwidth of up to 512 GBps.

Intel Stratix 10 MX Device Key Benefits

Intel Stratix 10 MX devices provide key benefits to address the challenges faced by next-generation system designers.

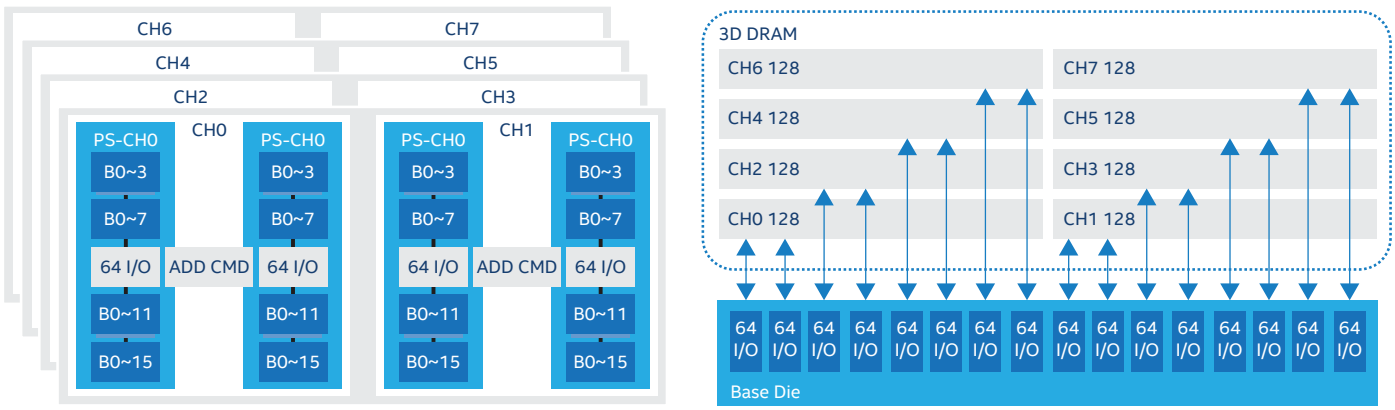


Figure 8. Logical Representation of Four High Samsung HBM2 Device with 16 Channels

Lower Power

Intel Stratix 10 MX devices provide lower power compared to conventional DDR solutions. For example, to meet 128 GBps of memory bandwidth, a system designer would need roughly 5 DDR4 (3,200 Mbps, 4 Gb) DIMM devices, with each consuming around an estimated 4 W (I/O + PHY + controller + memory) for a total estimated power of 22 W (assuming 30% read, 30% write, single rank configuration).† A single Intel Stratix 10 MX device provides equivalent memory bandwidth at approximately half the power. See Figure 10. This power savings is due to lower I/O power (short trace to EMIB vs. long PCB trace), lower data rates, and 3D DRAM stacking efficiencies. With no termination, the I/O buffer has lower overall capacitance which reduces I/O current consumption.

Smaller Form Factor, Reduced Board Complexity, and Ease of Use

As mentioned earlier, Intel Stratix 10 MX devices provide up to 512 GBps of bandwidth in a single package, resulting in substantial form factor savings. Figure 11 plots estimated form factor savings relative to a DDR4 3,200 Mbps DIMM (133 mm x 30 mm) for a range of memory bandwidth requirements. A single Intel Stratix 10 MX device provides equivalent bandwidth at a fraction of the form factor size, resulting in 15X form factor savings on average.†

This form factor savings frees critical board space, which gives designers flexibility to add more functionality or make better system-level performance or power trade off decisions. Eliminating DDR4 PCB routing results in lower board complexity (routing and layer count), improved signal and power integrity, and lower BOM and assembly cost. Additionally, the Intel Stratix 10 MX device solution is significantly easier to use.

High bandwidth, low power, and small form factor make Intel Stratix 10 MX devices attractive for a variety of end applications and markets.

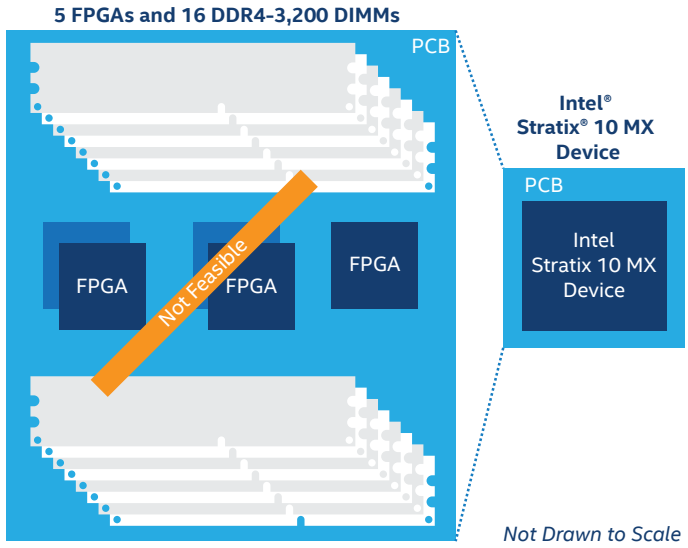


Figure 9. Comparing 400 GBps Memory Bandwidth Implementation

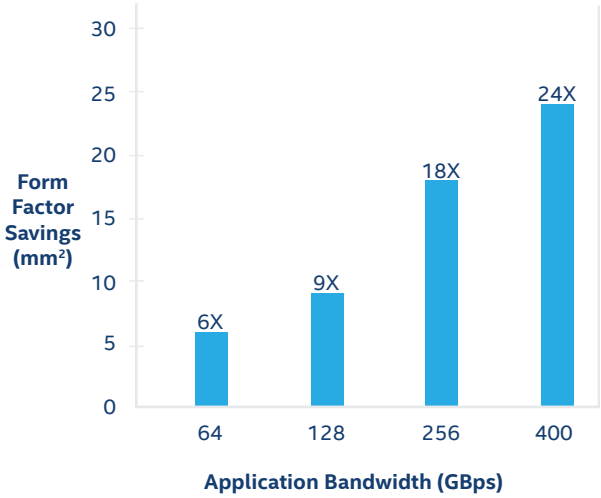


Figure 11. Estimated Form Factor Savings across Memory Bandwidth Requirements

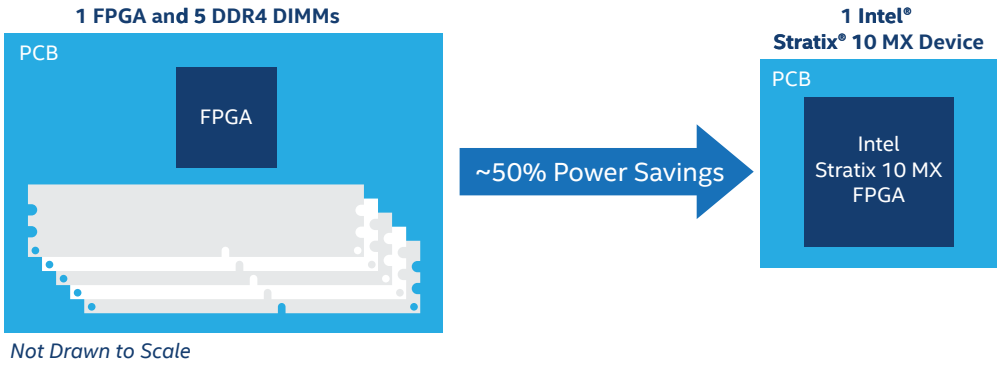


Figure 10. Example Power and Form Factor Savings (Application with 128 GBps Bandwidth)

Data Center Applications

In data centers, FPGAs add value by enabling new data touch and analysis applications. Combining Samsung HBM2 with a high-performance, monolithic FPGA fabric increases DRAM bandwidth to the FPGA by an order of magnitude, and lowers power compared to conventional DDR solutions. This combination expands the potential applications.

New Layer in the Memory Hierarchy

Offload engines with dedicated local memory can process unique workloads. The local memory enables an offload engine to perform memory intensive tasks without impacting the CPU RAM access. Samsung HBM2 DRAM enables new capabilities for offload engines in terms of speed and capacity.

Memory access is a trade-off between capacity and speed (bandwidth or latency). FPGAs have traditionally had 2-layer memory. The FPGA fabric has distributed RAM blocks, providing large amounts of bandwidth and concurrency with small capacity. The FPGA can connect to DRAM, giving much larger capacity for less bandwidth and greater latency. Figure 12 shows an FPGA in the data center before High Bandwidth Memory.

A Intel Stratix 10 MX solution enables a new point in the bandwidth/capacity curve. Adding the Samsung HBM2 memory to the overall solution spans the gap in the FPGA memory hierarchy.

Application Impact

Addressing this FPGA memory gap unlocks new applications. For example, dictionary searches and comparing pre-computed intermediate field search results reach new levels of performance. FPGAs excel at building high-touch data handling offload functions in diverse fields such as deep packet inspection, search acceleration, and security.

Samsung HBM2 also enables Intel Stratix 10 MX devices to reach new heights in DRAM access concurrency. Each High Bandwidth Memory interface supports 16 channels, therefore, a single package can scale up to 64 DRAM channels, a significant increase from 4 to 6 channels of external DRAM. With greater access concurrency, data center solutions (such as a table look up accelerator) can scale to a greater number of threads.

One of the key functions for FPGA offload involves data extraction and comparison with in-memory data structures. For these access patterns, the increased bandwidth, channel count (from 4 interfaces to 64 channels), and increased open bank count (from 64 banks to 512 banks) have positive impacts on memory subsystem performance. The extra channel count and bank pool count allows more accesses to hit successfully on open DRAM banks. Because operations can avoid the bank activation penalty, this implementation increases performance. Data center applications process a large number of threads in parallel, therefore, this implementation provides a strong advantage. Additionally, the open bank accesses enable lower memory power by avoiding unnecessary power hungry bank activates and pre-charges. For key tables, DRAM bank and port counts are

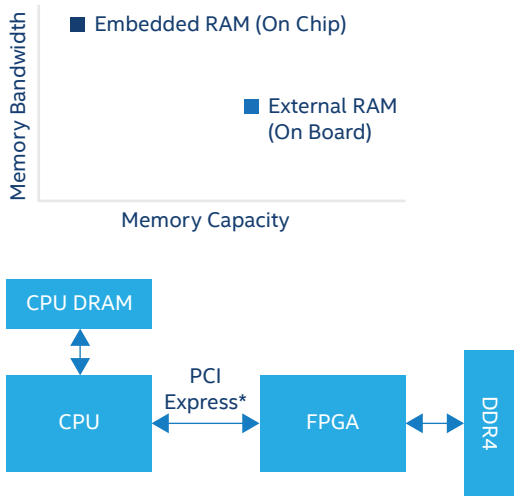


Figure 12. Traditional FPGA Data Center Applications

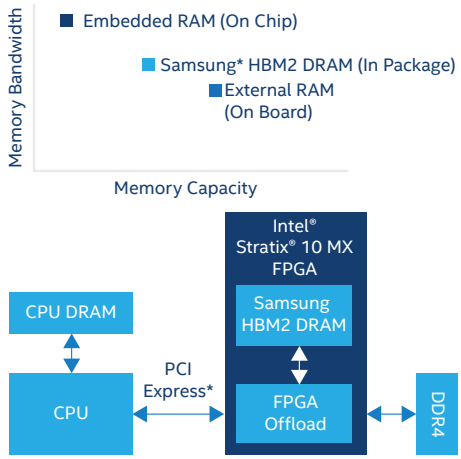


Figure 13. Intel Stratix 10 MX Data Center Applications

high enough that duplication can reduce effective access time.

The overall FPGA memory structure enables unique flexibility because the memory subsystem can be built around an application instead of the application meeting the fixed hardware of the memory subsystem. It is also easy to customize memory controller policies in an FPGA system.

Finally, the Intel Stratix 10 MX device solution enables intermediate sized data structure storage (in Samsung HBM2) and larger data structure storage in DDR DRAM. This capability is unique to the scalable memory subsystems that designers can build with Intel Stratix 10 MX devices.

Application Summary

Intel Stratix 10 MX devices allow greater bandwidth, greater access concurrency, and more open bank accesses than conventional DRAM solutions. These attributes create a DRAM memory subsystem for intermediate-sized tables for accelerator solutions. Additionally, designers can combine the integrated HBM DRAM with block RAM and DDR DRAM to build a comprehensive memory solution.

Algorithmic Acceleration for Streaming Cyber Security Analytics

For high data rate (greater than 10 GbE) streaming cyber security analytics applications, using FPGAs to accelerate algorithms is challenging due to the lack of abundant, high-throughput local memory. Current architectures use large DDR3 or DDR4 memories to store interesting data. Expensive memory reads and writes to DDR3 or DDR4 memory and the limited size and memory bandwidth make it difficult to coordinate data on chip and extremely difficult to access data off chip in an efficient manner.

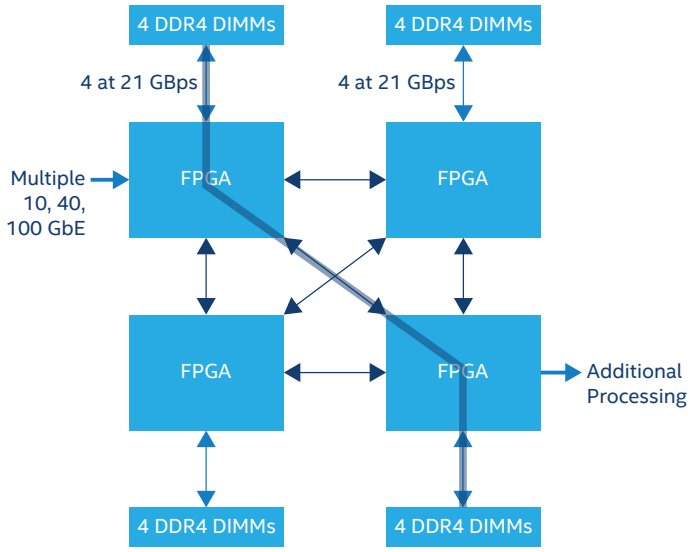
Intel Stratix 10 MX devices address streaming cyber security analytics critical need with high on-chip bandwidth to coordinate between multiple high-data-rate streams and allow relatively large on-chip storage. The reduced latency allows applications to search multiple 100, 400, and 1,000 GbE streams and coordinate them within a smaller compute node. Intel Stratix 10 MX devices lower the FPGA count for each compute node because each FPGA has greater storage and higher on-chip bandwidth.

For many high-data-rate streaming applications, maintaining data movement is a large, and growing, problem. As an example, current applications must fan out data streams to multiple CPUs or FPGAs to perform the security analytics, which introduces latency from the complex external fabric and inter-chip timing. When even small amounts of data must be stored off chip in this more complex architecture, accessing this off chip memory is expensive due to latency, increased power for additional DRAMs, additional computing elements, and overall lower PFLOPspsq ft. This slowdown in data movement rates constrains the system architecture to lower overall data bandwidth to compensate for slower memories.

Figure 14 shows four FPGAs with 128 GB of DDR4-2666 memory with an estimated 336 GBps memory bandwidth. Figure 15 shows the Intel Stratix 10 MX device with 1,024 GBps bandwidth and approximately 20% power consumption and 20% PCB area compared to Figure 14.† For the same power budget and PCB floorplan, an application could implement four Intel Stratix 10 MX devices for the same sub-system. With the available transceivers, a fully meshed subsystem can handle many 512 GBps Samsung Ethernet or Infiniband streams that would have otherwise required many racks. The increased meshing leads to lower latency between elements to create a larger database.

In Figure 14, the correlated data latency across different DDR4 DIMMs associated with fully meshed FPGAs is considerable, even on well-designed systems. The shaded line shows the system architect's worst-case latency with many milliseconds delay.

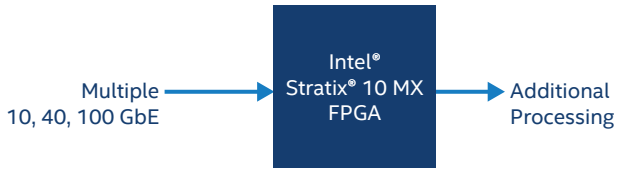
In addition, in Figure 14, when the application needs to perform upstream feedback or downstream forwarding of the database, the coherently combining the 16 DDR4 DIMMs is much more difficult than the Intel Stratix 10 MX device example. A traditional approach keeps multiple copies of the database to reduce latency, which requires database management to maintain coherency before sending upstream or downstream data. The Intel Stratix 10 MX device design eliminates coherency problems and simplifies the design. Additionally, the application maintains data movement and eliminates the database management overhead.



Current Streaming Applications
- Total Memory Bandwidth = 16 x 21 GBps = 336 GBps
- Complex Serializer/Deserializer (SERDES) Mesh Design
- Complex Database Break Up over 16 DIMMs
- Forced Data Partitioning

Figure 14. Traditional FPGA Partitioning for Streaming Cyber Security Analytics

Significant Bandwidth, Power, and Form Factor Gains vs. Traditional Methods



New Streaming Applications
- Total Memory Bandwidth = 1 TBps
- 3X Memory Bandwidth
- Simplified Design
- Unified Database
- Maintain Data Momentum

Figure 15. Improved Architecture with Intel Stratix 10 MX Device

Conclusion

Memory bandwidth is projected to grow dramatically to meet next-generation system requirements. A variety of end markets and applications (such as data centers, high-performance computing (HPC), broadcast, data analytics, wireline networking) drive this memory bandwidth growth. These next-generation systems demand the highest levels of memory bandwidth, the lowest power, and smallest form factor.

Conventional memory systems such as DDR3, DDR4, QDR, and RLDRAM are struggling to keep pace with this explosion in memory bandwidth. These conventional solutions cannot simultaneously meet the key requirements: higher memory bandwidth, lower power, and smaller form factor. Emerging memory technologies using 3D stacking technology like Samsung HBM2 can satisfy this bandwidth requirement.

Intel Stratix 10 MX devices are a new class of product that uses Intel's patented EMIB technology to efficiently integrate Samsung HBM2 with a high-performance monolithic FPGA fabric in a single package. Intel Stratix 10 MX devices provide up to 512 GBps of aggregate memory bandwidth in a single package. The Intel Stratix 10 MX device solution is specifically designed to meet high-performance system demands where bandwidth is paramount, providing 10X higher memory bandwidth, lower power, and smaller form factor compared to conventional solutions.†

References

[2] http://www.intel.com/content/www/us/en/foundry/emib.html?wapkw=emib
[3] http://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/wp/wp-01251-enabling-nextgen-with-3d-system-in-package.pdf
[4] http://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/wp/wp-01258-achieving-highest-levels-of-integration-in-programmable-logic.pdf

Where to Find More Information

For more information about Intel and Intel Stratix 10 MX FPGAs, visit https://www.altera.com/products/sip/memory/stratix-10-mx/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.