Chiplet Interconnect – Parallel or Serial?

By Michael Miller
Chief Technology Officer
MoSys, Inc.

What is a Chiplet?

Chiplets are the latest trend in the industry. By Wiki's definition, "a chiplet is an integrated circuit block that has been specifically designed to work with other similar chiplets to form larger more complex chips. In such chips, a system is subdivided into functional circuit blocks, called 'chiplets', that are often made of reusable IP blocks." By this definition, nearly everything that was built to go on a circuit board or the classic multi-chip module over the last 20 years could be considered a chiplet. The key difference today is that these chiplets could be in die form (like multi-chip modules) but connected into extremely dense systems using interposer technology (basically connecting with silicon substrates vs classic substrate material for much higher pin density).

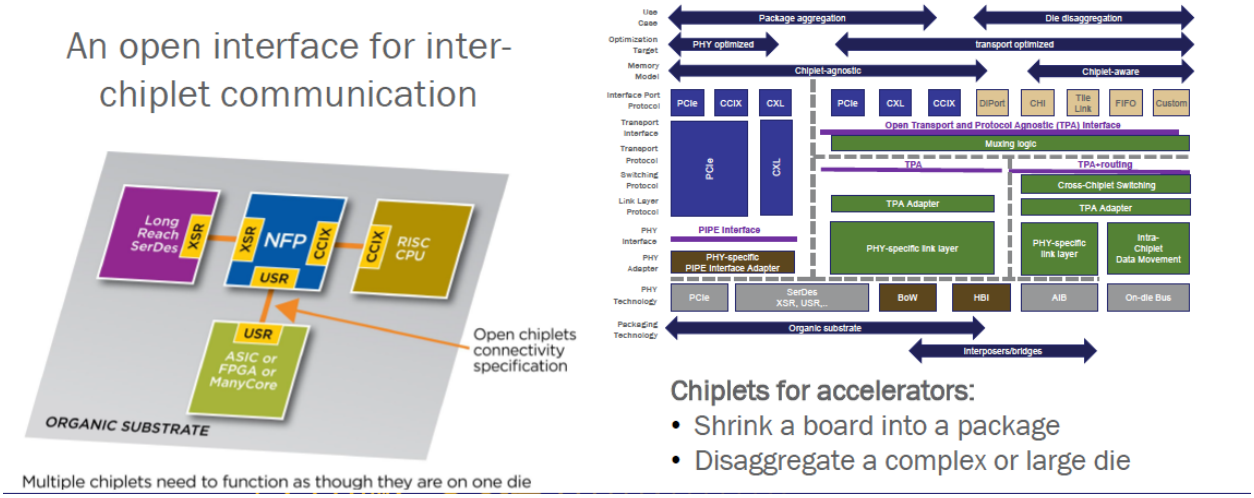


Image source: OCP ODSA

Several critical factors are driving these trends and they are the usual – cost, size and time to market:

- Cost: The cost of IC development is increasing dramatically. This includes the cost to develop (tools, time, complexity), the cost of the materials and the cost of the mask sets. All of these are growing exponentially with current estimates to generate a new IC in excess of $100M for the latest FinFet technologies and this does not even include the software. For these kind of expenses, significant volume is generally needed to get an acceptable return on the investment.
- Size: Size can affect the decision in several ways. First, as the chip gets larger, yields drop exponentially, so costs go up. Second, there is a desire to fit more in the same size package – generally 2-4x more performance in same footprint. Chiplets can break the die into higher yielding parts, will allow small footprints and will allow heterogeneous technologies to be packed close together.
- Time to market: Recreating mixed signal capabilities in each node is incredibly time consuming, expensive and risky. Chiplets allow for connection of heterogeneous technologies. In addition, instead of creating many individual designs of different sizes, chiplets may allow for multiple sizes, and versions of a technology without having to design another piece of silicon.

Note that the other usual critical attribute is power but any I/O increase generally has a negative impact on power, thus power becomes a trade-off for chiplets.

Industry Efforts to Standardize

The industry is keenly interested in chiplets and government and industry bodies have been initiated to create solutions: DARPA created a program called CHIPS and an industry group called ODSA (Open Domain-Specific Architecture) which is now under the OCP (Open Compute Project) umbrella is working to establish industry standards and an ecosystem for chiplets. For more information on these standards you can check out the following direct links: DARPA CHIPS program and ODSA.

The list of companies using the technologies is large and growing. Different forms of chiplet technology have been used in Field Programmable Gate Arrays or FPGAs, compute technology and networking for connecting memory and/or other heterogeneous elements and by networking companies building massive switch matrix implementations. The challenge is that most of these examples have been connecting technologies built and controlled by a single company. This is quite limiting relative to development time, cost and flexibility. It also requires highly vertical integration and scale.

The most visible example of a high pin count, interposer chiplet concept from an industry perspective is HBM or High Bandwidth memory. This stacked DRAM technology is built by multiple memory vendors and is connected to industry devices through a standard high pin count parallel interface. The development of this technology was driven by memory bandwidth, memory size and footprint in server/compute applications but it has seen expanded application areas. Interfaces leveraging HBM type interfaces have been the primary focus for the parallel options for chiplet interconnect.

Key Decision for Chiplets is the Interconnect

The key decision regarding chiplets is the interconnect or I/O. There are two basic options available – serial or parallel. Fundamentally, think of this as narrow and fast (fewer lanes running at high data rates) or wide and slower (100s or 1000s of lanes running at slower rates). To put this in perspective, MoSys Bandwdith Engine family of serial memories have a SerDes interface that is 16 lanes (64 total pins) of differential I/O that run at up to 28Gbps (800Gbps of I/O). An alternative to that same bandwidth would be a parallel interface that would require 800 lanes (800 pins) of parallel I/O running at 1GHz per pin to deliver the same bandwidth. While latest interposer solutions have made huge pin counts feasible, that is still a lot of pins. Each interconnect methodology has its trade-offs. When selecting interconnect, one size does not fit all. We will explore this in more detail below.

Parallel Chiplet Interconnect

Parallel busses such as the High Bandwidth Interconnect used in the High Bandwidth Memory deliver high bandwidth, low latency and relatively low IO power. In general, the parallel interface will deliver the highest bandwidth and lowest latency at the lowest power. Prior to interposer technology, the parallel interface was just too wide, making chip sizes large and high bandwidth interconnect unfeasible. Interposer technology has provided a solution to this limitation by drastically increasing the pin density between chips. The interface allows for massive pins and bandwidth but creates chip layout constraints and forces "lego block" like form factors where every chip needs to be of fixed aspect ratios and multiples of a fundamental width. This is not the natural way of the world. Just examine any PCB – devices and underlying silicon chips tend to come in all shapes and sizes. Forcing devices to specific sizes just to fit the aspect ratio works against fundamental silicon economics.

In addition, the assembly of these massive parallel interfaces using interposers (IC interconnect substrates) make assembly more expensive and managing power proximity challenging. In the case of HBM, placing a memory whose performance is inversely proportional to temperature right next to a 100-watt processor or ASIC creates thermal challenges. In the end, if power and size are highest priority, parallel connection may be the best alternative.

Serial Interconnect

Serial connections have been used in the industry for decades. These have been used for long reach interfaces across the network or chip to chip interconnect with protocols such as Interlocken and PCIe across PCB boards. The interfaces are high bandwidth, pin-efficient, proven and extremely flexible. Serial connections can address lots of the proximity problems of parallel interfaces such as: chip offsets, routing around comers, pitch mismatch and differing chip aspect ratios. The lower power Very Short Reach (VSR), USR (Ultra Short Reach) or Extra Short Reach (XSR) chip to chip interfaces

allow for low cost connectivity across a standard substrate and flexibility to separate devices across a standard substrate for convenient placement.

The typical drawbacks brought up for serial interfaces are power and latency. The power of high-speed serial interfaces are generally higher than single ended parallel connections. Added latency comes from the fundamental serialization and deserialization that is required with the high-speed interfaces and from the serial protocol used.

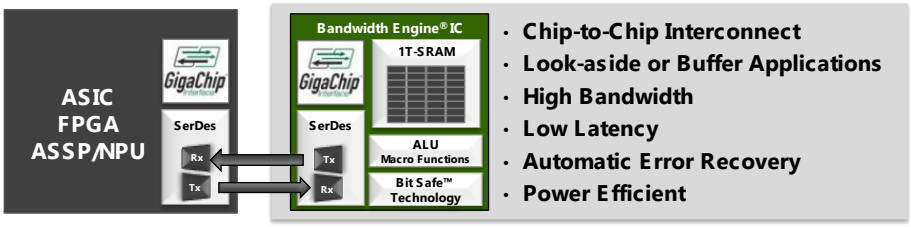
Serial Connections Provide a Viable and Proven Chiplet Interconnect

While the parallel interface is garnering much attention for chiplet interconnect due to the proliferation of HBM with its implied power and latency benefits, serial options should not be disregarded. Many of the trade-offs are not as significant as the myths would suggest.

Bandwidth density and power are a big component of discussion. Power fundamentally comes down to voltage swing, parasitic capacitance and frequency with voltage swing playing an exponential contributing role (V^2 x C x Freq). Every signaling whether it is single ended as in parallel busses or differential as in SerDes must follow this rule. New standards for serial interfaces like USR and XSR trade voltage swing for distance. The 28G SerDes interfaces on the MoSys bandwidth engine consume less than 5-10% of the total chip power and this includes provisions for connectivity over mezzanine cards. The power could be reduced significantly by making the trade-off above. Another example of low power SerDes chip to chip interconnect is: Achronix shows possible chiplet solutions using SerDes.

Serialization also need not add a lot of latency when using 8 to 1 muxing such as in DRAM. For example, MoSys products incorporate a CEI-25G SerDes where the total Tx + RX including deskew latency is under 3ns. This is well within the goal of ODSA/DARPA. In addition, protocol latency should not be disregarded. It can also be reduced significantly. In memory products such as the MoSys Bandwidth Engine, latency is critical. The MoSys Gigachip Interface protocol provides nearly an order of

magnitude lower latency than other serial protocols to help minimize latency. By combining both the SerDes and protocol latency reductions, the latency penalty of serialization can be reduced significantly.



The serial interface designed for the MoSys Bandwidth Engine family encompasses the critical power and latency attributes for a chiplet. The only real difference today is that the Bandwidth Engine devices are delivered on substrate material instead of a die.

What Serial Protocol is best - NRZ, PAM4 or other?

The current trend for high speed physical layer devices (PHYs) is to use PAM4 signaling when higher rates and longer reach are needed. Networking standards which tend to drive PHY standardization have primarily transitioned to PAM4 at 50Gbps to enable higher data throughput, as PAM4 can deliver twice the bit rate for a given baud rate vs traditional NRZ. But it is also more complex and higher power and more likely to require forward error correction and complex DSP functions to maintain acceptable bit error rates. Forward Error Correction, if required, introduces asynchronous latency that can be as high as hundreds of ns. That kind of additional latency is likely that make performance sensitive chiplet interfaces intractable. NRZ provides the lowest power, lowest latency serial interface option for the short reaches (in tens of mm or less) used for chiplets. Other non-standard options such as the Chord signaling from Kandou may use a mix of NRZ and signaling techniques.

A key benefit for using the PAM4 interface is that it is needed for the long reach high bandwidth interconnects. Many members of the industry would prefer to consolidate PHY efforts towards a single protocol type, but it is unlikely one interface will efficiently

serve all these applications. In general, an industry standard NRZ interface is the ideal option for serial chip to chip interconnect until the total bandwidth per pin mandates the use of more complex signaling options. NRZ provides the lowest latency and lowest power. With silicon examples of 56G and 112G NRZ already shown for short reach, it could be quite some time before a more complex PAM4 interface would be required.

Conclusion

Making a choice for Chiplet interconnect is not obvious and the number of different options currently in use, proposed as part of the standards process and in discussions at ODSA tell us that. While there is much excitement around parallel interfaces, serial technology provides a proven, scalable and flexible interface that allows for low cost manufacturing and packaging. The power and latency drawbacks may not be as dramatic as parallel proponents would suggest.

As part of the Bandwidth Engine technology, MoSys has built a serial interface that is both low latency and low power and is relevant to chiplets at a fundamental technology level. The PHY interface follows industry standards and the low latency GCI interface definition is built for low latency transfer of any type of data from chip to chip. If manufactured in die form, the Bandwidth Engine itself is a virtual chiplet. By leveraging existing low latency serial technologies like those developed for the MoSys Bandwidth Engine family, the industry has viable serial options that will deliver higher flexibility, lower cost options using standard substrate materials while minimizing the power and latency penalties.

###