White Paper

Using 10-Gbps Transceivers in 40G/100G Applications

This white paper identifies the key drivers behind the migration to 100G interfaces, and shows how to leverage the unique ability of FPGAs to implement this high-speed interface. The emerging 40GbE and 100GbE standards for data center and core network systems rely heavily on FPGAs to share those sectors with other protocol infrastructures. In addition to providing an unprecedented amount of resources such as logic, on-chip memory, and DSP blocks, Stratix IV devices are the only FPGA family to enable these 40G/100G designs, which require 10G transceiver data rates with extraordinarily low jitter performance to meet high-speed design requirements.

Introduction

Using recent technological advancements, the current generation of FPGAs has the bandwidth, high transceiver count, and ability to support multiple protocol standards on a single device. Telecommunication equipment manufacturers—focused on developing the next generation in bridging applications and switching solutions for 40- and 100-Gigabit Ethernet (GbE)—are the emerging target market segment for FPGAs, which can meet the high speed data rates and bandwidth requirements.

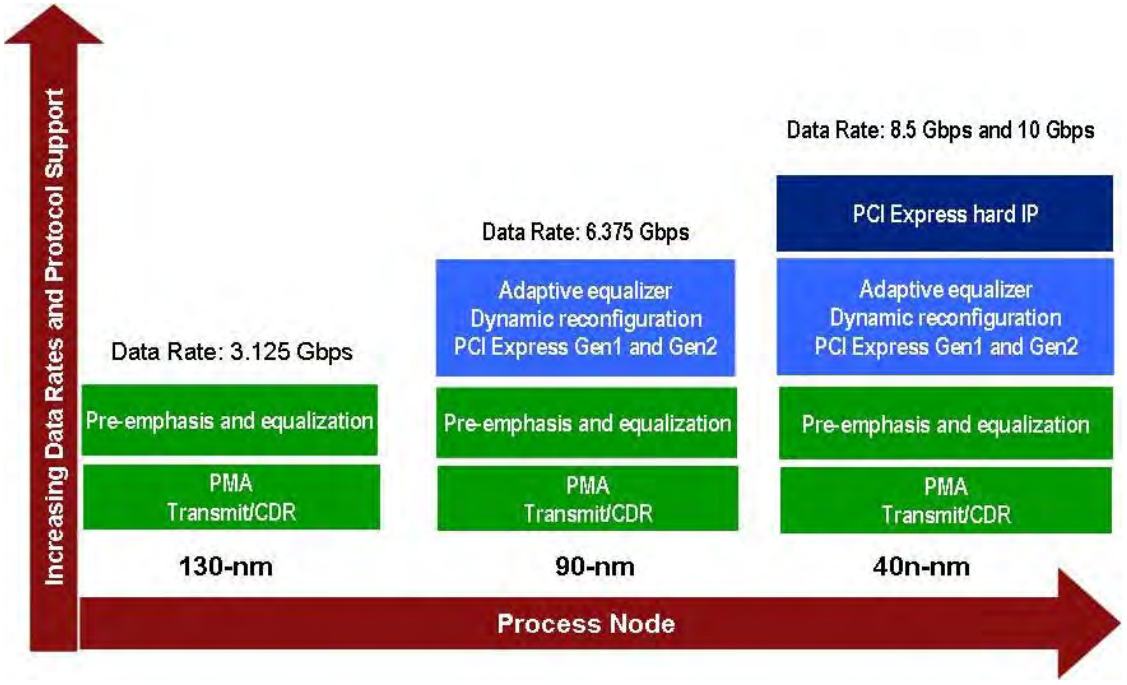
This white paper identifies the key drivers behind the migration to 100G interfaces, and shows how to leverage the unique ability of FPGAs to implement this high-speed interface. Standards evolution for this protocol is crucial to defining common implementation, which ultimately provides the means for cost efficiencies through economies of scale in production of key cost components. Consolidation to common interface standards facilitates the simplicity of architecture and operations. Ethernet interfaces are well defined for 10/100/1000 Mbps and 10 Gbps, with standards currently in the draft phase for moving beyond 10 Gbps.

The emerging 40GbE and 100GbE standards for data center and core network systems rely heavily on FPGAs to share those sectors with other protocol infrastructures. (These include Fibre Channel, Infiniband, and SONET for more bridging and data aggregation types of application.) As LAN speeds increased to a gigabit, the most cost-efficient networks embraced Ethernet as the primary data link protocol, which could be satisfied by ASSPs and FPGAs. But as LAN, SAN, and MAN speeds surpass 10G, the most cost-efficient networks are embracing multiple data link protocols, which must use multiple ASSPs or deploy FPGAs as bridging devices to provide cost-effective solutions.

Meeting the 40G/100G Requirements

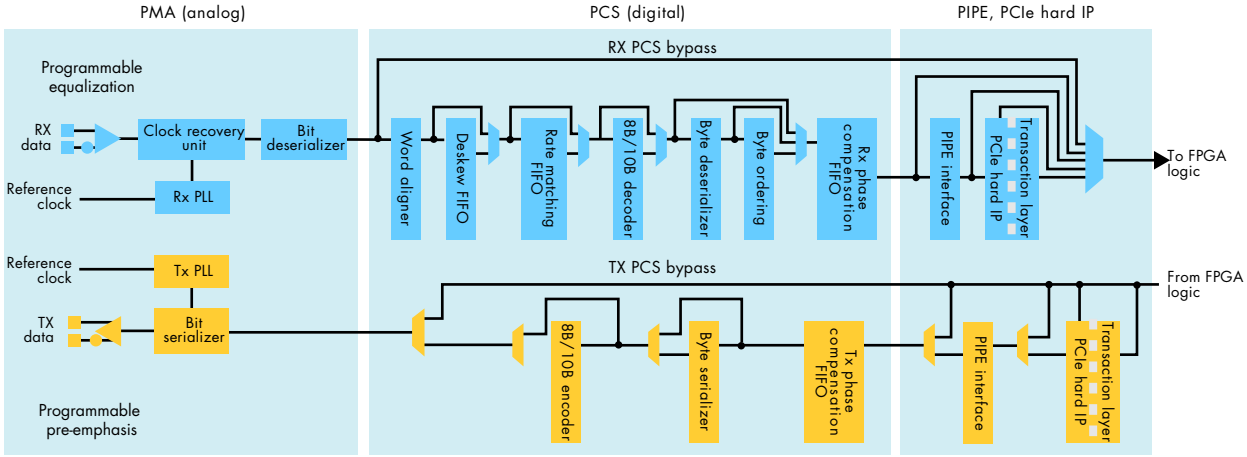
To address the 40G/100G requirements, FPGAs have evolved over the last few process generations, as shown in Figure 1. At the 130-nm process nodes, FPGAs only supported up to 3.125 Gbps, while today's 40-nm process nodes support data rates beyond 10 Gbps.

Figure 1. Process Node Generations



Altera® Stratix® IV GX FPGAs have up to 32 embedded transceivers supporting data rates from 600 Mbps to 8.5 Gbps, plus an additional 16 transceivers supporting data rates from 600 Mbps to 6.5 Gbps (a total of 48 transceivers up to 6.5 Gbps). The transceivers include both a physical coding sublayer (PCS) and a physical media attachment sublayer (PMA), the two of which enable Stratix IV FPGAs to implement standard and proprietary protocols. Figure 2 shows a diagram of the Stratix IV GX transceiver block.

Figure 2. Stratix IV GX Transceiver Block



Moving to smaller process geometries allows for greater system integration in a chip because the FPGA densities roughly double from one process generation to the next. While increasing the feature set of the FPGA is important, being able to simultaneously meet leading-edge system performance requirements while minimizing power is extremely important. Stratix IV FPGAs leverage proven processing techniques and architectural innovations like Programmable Power Technology and dynamic on-chip termination to minimize system power while supporting high system bandwidths of 40G and 100G applications.

For detailed information on the process and architectural innovations that enable Stratix IV FPGAs to lower power consumption, refer to Altera's *40-nm FPGA Power Management and Advantage* white paper.

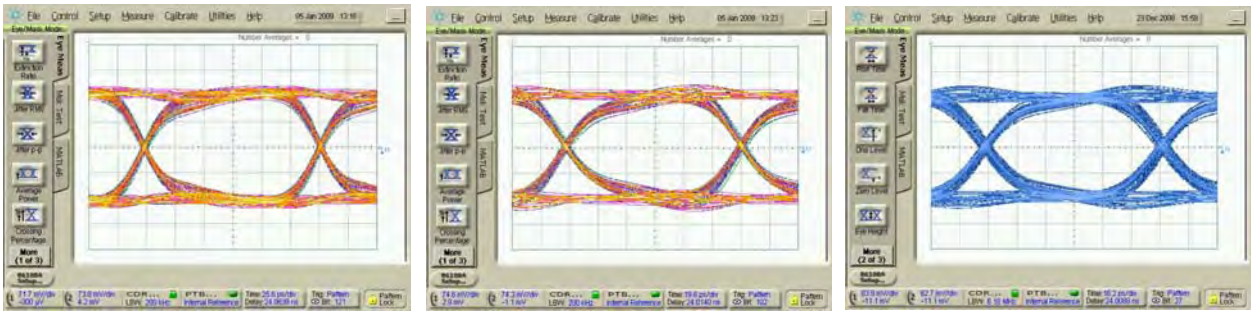
The PMA is an embedded macro dedicated to receiving and transmitting off-chip high-speed serial data streams. The PMA channel consists of full duplex paths (transmit and receive) with I/O buffers, programmable output voltage, pre-emphasis and equalization, clock data recovery (CDR), and serializer/deserializer (SERDES) blocks. Table 1 shows the transceiver power (PMA only) per channel for Altera's 40-nm transceiver products.

Table 1. Transceiver Power per Channel (PMA Only)

Data Rate (Gbps)	40-nm Transceiver Power (mW)
3.125	100
6.5	135
8.5	165
10.3	190
11.3	200

Stratix IV GX transceivers use advanced power supply regulation and filtering techniques to reduce transmitter jitter and improve receiver jitter tolerance. This allows the transceivers to exhibit superior bit error rate (BER) performance when used in real system links. On-chip voltage regulators for both transmit and receive phase-locked loops (PLLs), careful isolation of sensitive analog circuitry, and extensive use of on-die and on-package decoupling capacitors all contribute to a robust power distribution scheme for the transceivers while delivering exceptionally clean power to the analog circuits. Figure 3 shows the eye of the transmitter operating at 6.25 Gbps, 8.5 Gbps, and 10 Gbps.

Figure 3. Transmitter "Eye" at 6.25 Gbps (left), 8.5 Gbps (center), and 10 Gbps (right)



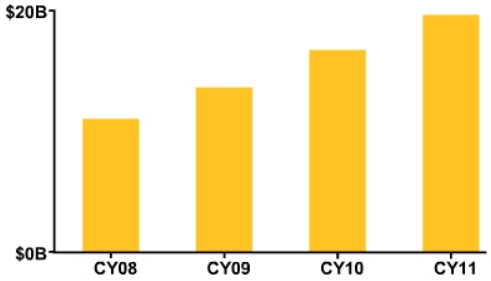
Demand for 100 Gbps

The rapid growth in network bandwidth is driven primarily by increasing numbers of broadband subscribers via xDSL, FTTx, WiMAX, and the 3G/4G platform. Network bandwidth applications such as IPTV, VoIP, and online gaming, as well as a growing number of online users accessing video-on-demand sites, also have influenced a demand for higher bandwidth. Today, dedicated networks, such as wireline carrier networks and cable operators, already coexist, providing broadcast and on-demand content, along with over-the-top video via the Internet. Independent of the source of content or the delivery mechanism, the network for such services is IP based with huge bandwidth demands in the core of the transport network.

Bandwidth is growing rapidly, and demand will only accelerate in the coming future (Figure 4). The current network infrastructure will scale to provide higher capacities by using the 10GbE protocol, but this is cumbersome from a data network-topology standpoint as each 10-Gbps interface requires a separate network instance. Link aggregation, specified in IEEE 802.3ad, is one approach that simplifies the data network topology by bonding multiple lower speed lanes, but it has limitations and is far from an optimal solution. There is no doubt that the requirement for higher speed interfaces will increase, as the 40-Gbps protocol is now being deployed in the core network, and applications are expected to accelerate. The protocol currently used is an aggregated form of single-lane 10G network

based on the 802.3ae protocol standard. Current solutions are based on SONET/SDH and Optical Transport Network (OTN) standards, but there is clearly a need for a 40GbE standard. 100 Gbps is also a key interest area, and the IEEE standards body has proposed an initial draft to define the 100GbE interface.

Figure 4. Demand for Bandwidth Is Increasing: 10G/40G/100G Worldwide Revenue



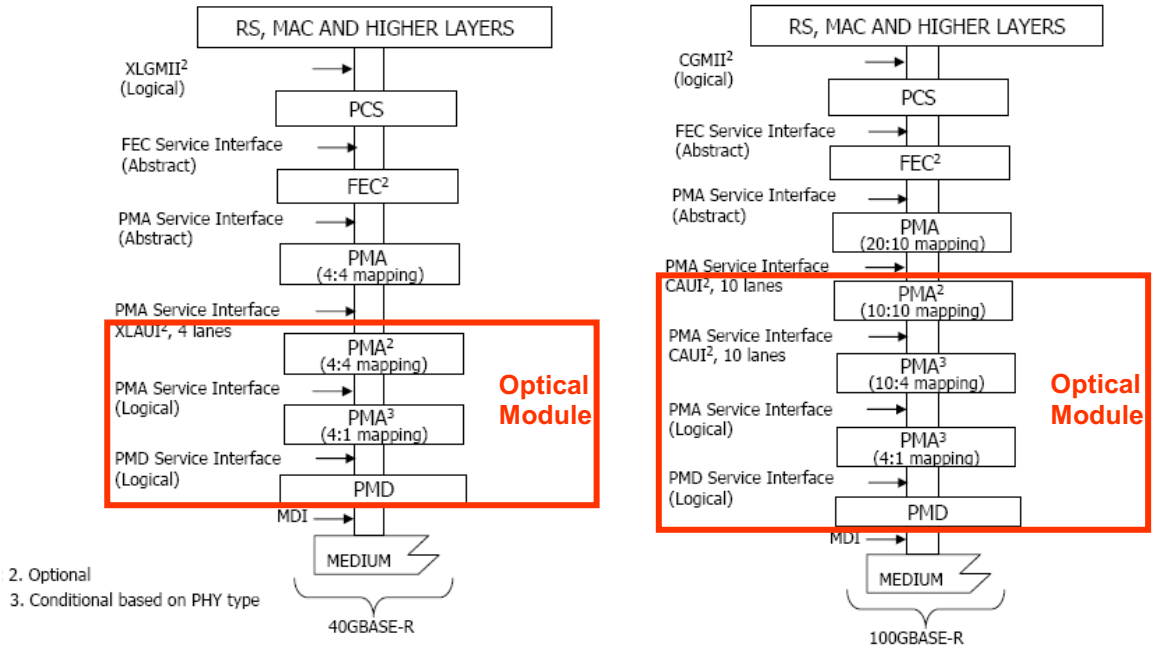
Standards Evolution for 40GbE/100GbE Interface

The IEEE 802.3 High-Speed Study Group (HSSG) formed in 2006 to study the market needs and definition of standards for a 100GbE interface protocol (shown in Figure 5). In 2007, an additional proposal to include a 40-Gbps rate was adopted. The proposal extends the 10G Base-R Carrier Sense Multiple Access with Collision Detection (CSMA/CD) standards to include both 40 Gbps and 100 Gbps with the following objectives:

- Include full-duplex operations only
- Preserve the 802.3/Ethernet frame format using the 802.3 MAC
- Preserve the frame size specification of the current 802.3 standard
- Support a BER better than or equal to $10E^{-12}$
- Comply with OTN for WAN applications

The initial draft from the HSSG, proposed on August 2008, provided the details of implementing this protocol. The FPGA market segment will play a vital role in evaluating the merits of the protocol and prototyping it on current platform.

Figure 5. 40GbE/100GbE Interface



40G/100G IP Solution for 802.3 Standards

The MAC and the PCS portion of the 802.3ba draft are implemented as soft logic on the FPGA fabric. As shown in Figure 6, the MAC feature is identical to the legacy Ethernet protocol except for the 100-Gbps data aggregation. The PCS implementation for 802.3 uses the 64/66 encoding scheme, which is identical to the 802.3ae standards.

Figure 6. MAC of 802.3ba Draft

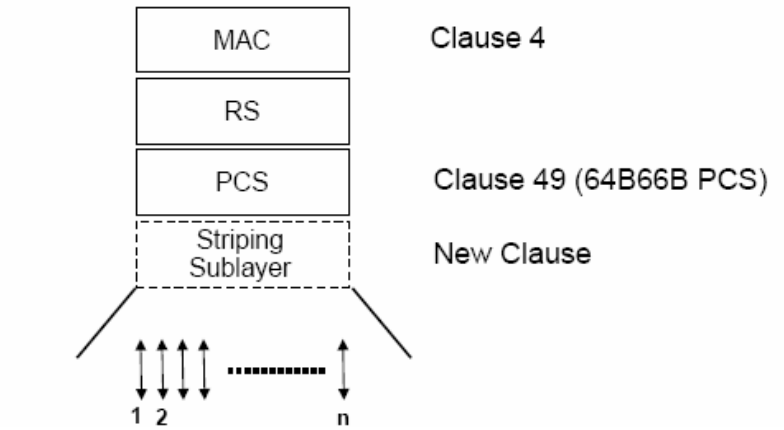


Figure 7 shows the transmit path of the PCS channel. Data from the MAC first is encoded into a continuous stream of 64B/66B blocks and scrambled. This 66-bit scrambled data then is stripped over 20 virtual lanes (VLs) via a simple round-robin mechanism. A unique marker (66B word) is added to each VL at the same time and on a periodic basis. These markers are used by the receiving PCS block to identify, deskew, and reorder the data from the VLs, and to re-assemble the 100G aggregate data stream. These 20 VLs will eventually multiplex to a 10-lane PMA, where each independent PMA lane runs at 10.3125 Gbps. The data is inverse-multiplexed across the 10 100G-10-bit Interface (CTBI) lanes. Overall, an in-band skew mechanism is implemented by the alignment markers within each VL.

Figure 7. Transmit Path of the PCS Channel

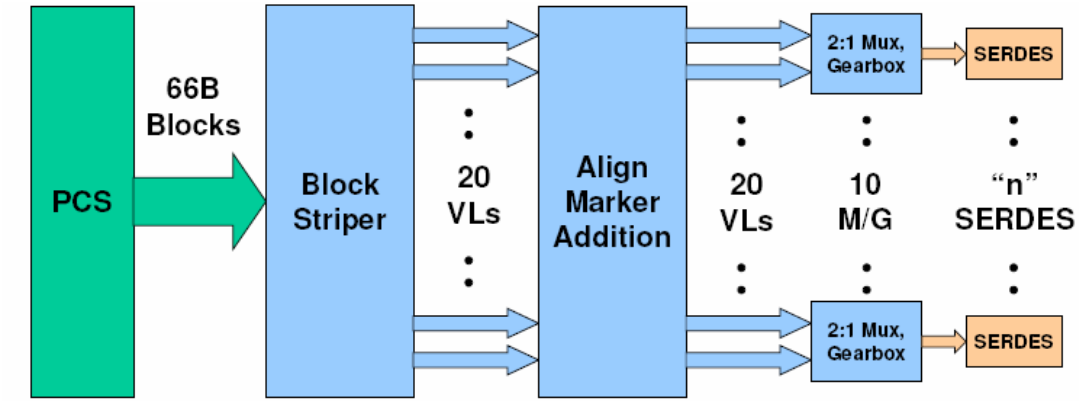


Figure 8 shows the receive path for the 100G PCS block. The receive channel demultiplexes the data from 10 (where *n*=10) electrical lanes to 20 VLs. The receive channel deskews and aligns the VLs to rebuild the 66B data stream. Each VL on the receive path has its own dedicated FIFO buffer, the overall depth of which determines the skew tolerance limit of the design. For example, a FIFO buffer that is two 64-bit words deep will tolerate a 128-bit skew.

Figure 9 shows the implementation of VLs, which are needed only when the number of electrical (*n*) and PMD (*m*) lanes are not equal. VLs make this transition easy when there is asymmetry in the device transceiver lanes. For the 100G implementation, the 100G aggregate stream is split into a number of VLs, which are based on 64B/66B blocks, and an alignment block is inserted within each VL. The number of VLs implemented is scaled to the least common

multiple (LCM) of the *n* and *m* lanes. A 100G implementation needs a maximum of 20 VLs and a 40G implementation needs a maximum of four VLs. The virtual lane marking allows the receive channel to perform skew compensation, realign the VLs, and reassemble the data stream to single a 100G or 40G aggregate (with the 64B/66B blocks in the correct order).

Figure 8. Receive Path for the 100G PCS Block

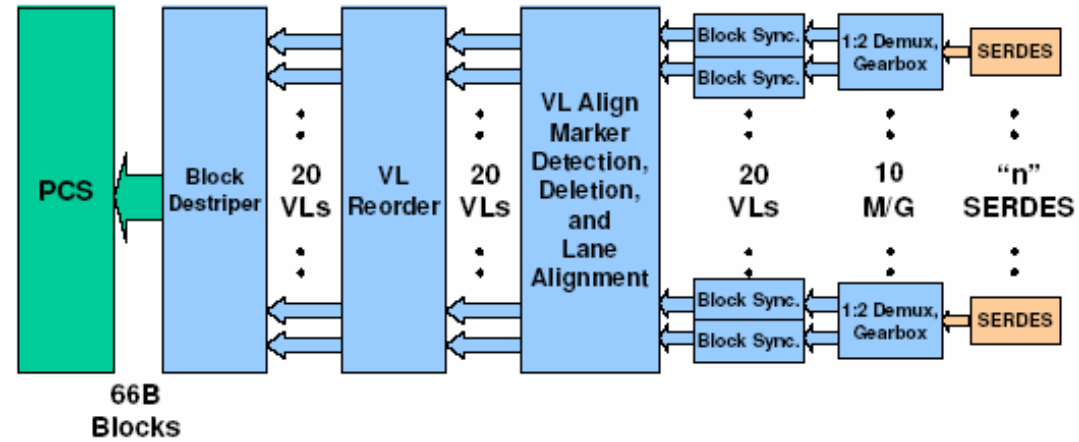
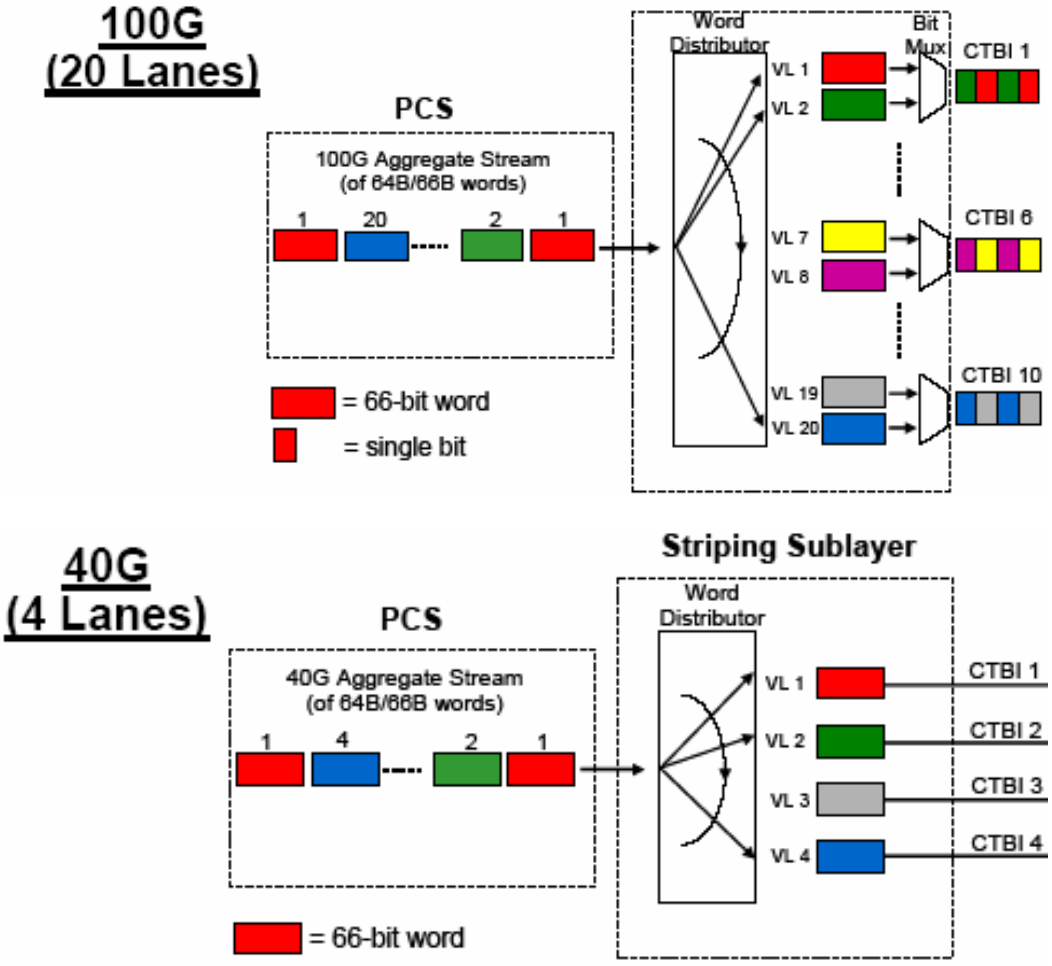


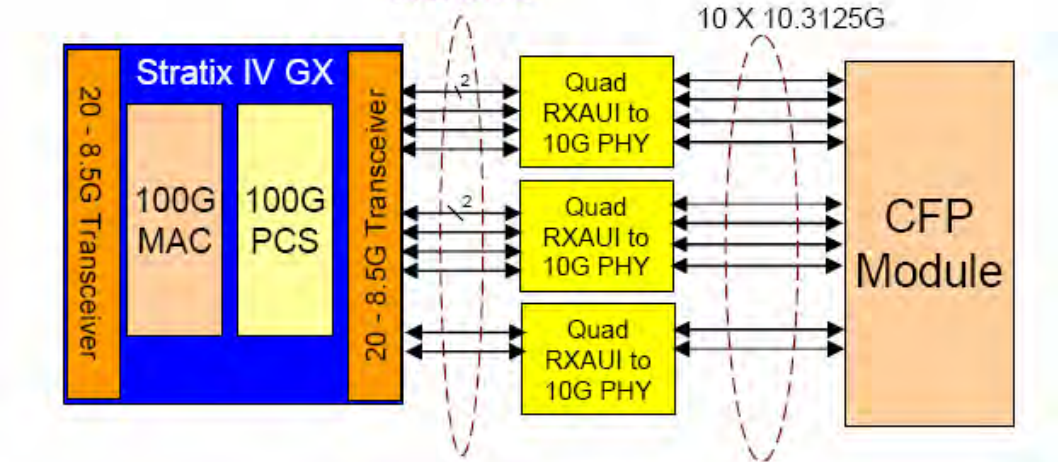
Figure 9. VL Concept



100G MAC Solution

As shown in Figure 10, Altera's 100G MAC solution (implementing 20 VLs) uses the Stratix IV GX FPGA along with an external ASSP or custom ASIC, which interfaces through RXAUI (a reduced XAUI interface running at double speed) and incorporates the MLD feature in the next-generation 10G PHY solution.

Figure 10. 100G MAC Solution

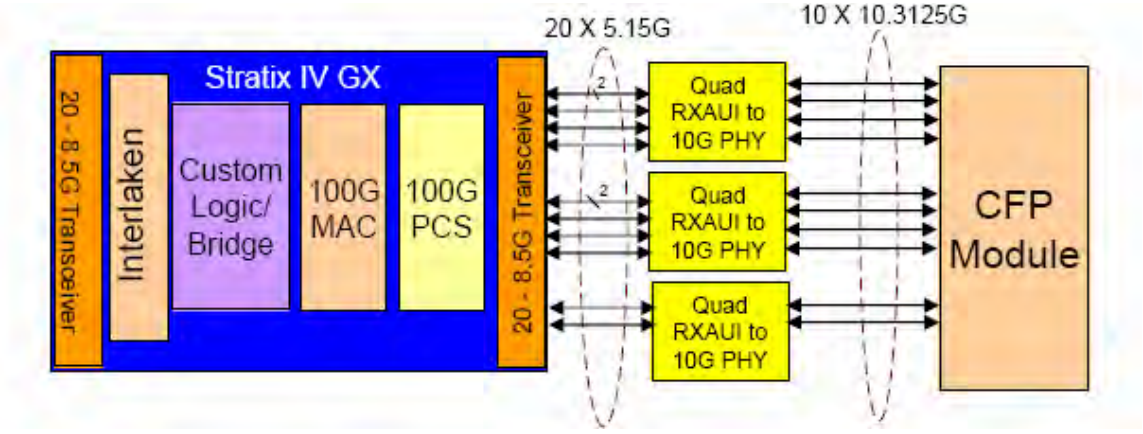


The 100G MAC IP is provided by Altera's partner MoreThanIP and is compliant with the 802.3ba standards. Altera is committed to work closely with MoreThanIP to provide a 100G solution.

Complete System Solution

On the system-side interface, the Stratix IV FPGA is capable of providing a 100G chip-to-chip interface using the Interlaken protocol via 20 lanes running at 6.375-Gbps data rate. As shown in Figure 11, the complete system solution provides a dedicated 100G data path from the line side to the system side. The flexibility of deploying an FPGA allows customers to leverage proprietary bus interfaces on the system side and to bridge between proprietary interfaces and the industry-standard Interlaken interface. Altera's Interlaken solution is compliant with the Interlaken Alliance specifications.

Figure 11. Complete System Solution



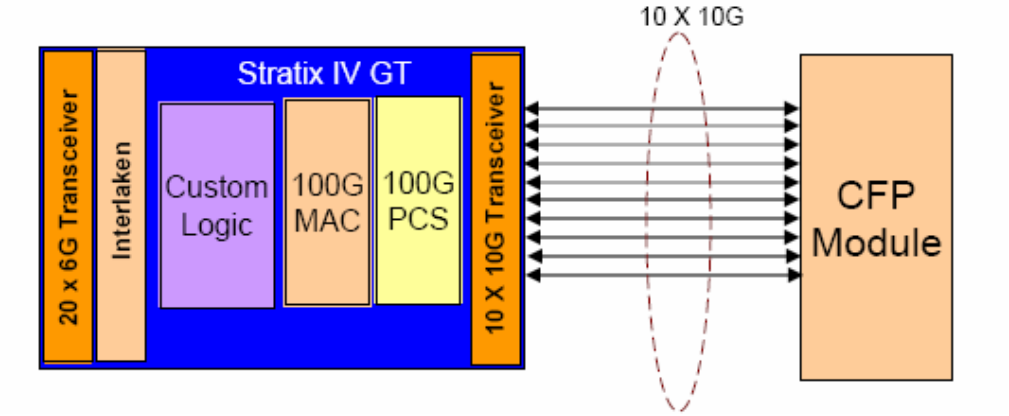
This proposed solution leverages the current ecosystem and third-party MLD-multiplexed PHYs to provide an industry-standard 100G platform at the 40-nm technology node. Altera is committed to providing optimal solution for implementing 100G on network platforms.

Please contact an Altera sales representative for further information on the 100G platform.

Optimized 100G System Solution

The optimized 100G system solution shown in Figure 12 leverages the transceiver technology leadership to operate beyond 10 Gbps, and eliminates the need for an external 10G PHY device capable of providing MLD features. The solution provides a true single-FPGA 100G solution, where the 100G CFP optical module can interface directly with the Altera Stratix IV GT FPGA. The Stratix IV GT device gives the customer the flexibility to incorporate features like traffic management, data policing, and packet processing on a single-FPGA fabric. In production today, the Stratix IV GT FPGA includes 32 transceivers, which are capable of running at 11.3 Gbps, plus an additional 16 transceivers (for a total of 48), which support a data rate of 6.5 Gbps.

Figure 12. Optimized 100G System Solution



Conclusion

The Stratix IV device family meets the market's demand for high density, high performance, and low power. It is the only FPGA family which provides an optimized 40G/100G solution. At the 40-nm process node, Stratix IV FPGAs provide an unprecedented amount of resources such as logic, on-chip memory, and DSP blocks. In addition, Stratix IV FPGAs enable 40G/100G designs, which require 10G transceiver performance with extraordinarily low jitter performance to meet high-speed design requirements. This device family is optimized to meet the demand for 40G/100G designs and is available today.

Further Information

- IEEE 802.3ba:
  http://grouper.ieee.org/groups/802/3/ba/index.html
- *40-nm FPGA Power Management and Advantage*:
  www.altera.com/literature/wp/wp-01059-stratix-iv-40nm-power-management.pdf
- *Innovating With a Full Spectrum of 40-nm FPGAs and ASICs With Transceivers*
  www.altera.com/literature/wp/wp-01078-stratix-iv-gt-40nm-transceivers.pdf
- Literature: Stratix IV Devices (E, GX, and GT variants):
  www.altera.com/literature/lit-stratix-iv.jsp
- Contact Altera:
  www.altera.com/corporate/contact/con-index.html

Acknowledgements

- Rishi Chugh, Product Marketing Manager, High-End Device Marketing, Altera Corporation
- Seyi Verma, Product Marketing Manager, High-End FPGAs, Altera Corporation

101 Innovation Drive
San Jose, CA 95134
www.altera.com