Introduction

Altera’s 40 nm Stratix® IV GX and Stratix IV GT devices provide a complete solution for developing 40 Gbps and 100 Gbps network line card applications. This application note describes how to implement a 40 Gbps/100 Gbps (40G/100G) Ethernet line-side interface using Stratix IV GT devices. It focuses on the following two implementation aspects:

- Transceiver configuration and clocking recommended for 40G/100G Ethernet
- Interfacing 40G/100G Ethernet Soft IP to the transceivers

Figure 1 shows an example 40G/100G network line card application with Stratix IV GT and Stratix IV GX devices.

Figure 1. Altera’s 40G/100G Network Line Card Application (1)

Note to Figure 1:
(1) Figure 1 shows the 40G/100G MAC, PCS, and PMA designed in a Stratix IV GT device with integrated 10 Gbps serial transceivers. You can implement this application in Altera's Stratix IV GX and HardCopy® IV GX devices with external RXAUI to XFI multiplexer chips between the Altera device and the optical module. For more information, refer to Using 10-Gbps Transceivers in 40G/100G Applications.

The three-chip 40G/100G line card solution shown in Figure 1 consists of:
Stratix IV GT Hard IP Blocks Used in 40G/100G Ethernet Implementation

The 40G Ethernet MAC implementation uses four full-duplex Stratix IV GT transceiver channels at 10.3125 Gbps, as shown in Figure 2.

Figure 2. 40G MAC Implementation in a Stratix IV GT Device
The 100G Ethernet MAC implementation uses 10 full-duplex Stratix IV GT transceiver channels at 10.3125 Gbps, as shown in Figure 3.

Figure 3. 100G MAC Implementation in a Stratix IV GT Device

Stratix IV GT Soft IP Blocks Used in 40G/100G Ethernet Implementation

Stratix IV GT devices do not have dedicated hard IP blocks that implement 40G/100G Ethernet Media Access Control (MAC), Physical Coding Sublayer (PCS), and gearbox logic. These blocks must be implemented with a soft IP core in the FPGA fabric.

40G/100G Ethernet soft IP cores targeted to Stratix IV GT devices are available from Altera’s IP partners MorethanIP (www.morethanip.com).

Physical Coding Sublayer (PCS) Block

In the egress direction, the PCS block is responsible for functions such as 64B/66B encoding, scrambling using an $X^{58} + X^{39} + 1$ polynomial, and striping the 66-bit data blocks into virtual lanes (VLs). In the ingress direction, it is responsible for functions such as 66B/64B decoding, descrambling using an $X^{58} + X^{39} + 1$ polynomial, and destriping the VLs into 66-bit data blocks. It also performs status monitoring and decoding functions as specified in the IEEE802.3ba specification.

Media Access Control (MAC) Layer

The MAC layer is responsible for functions such as Ethernet frame encapsulation and decapsulation, Frame Check Sequence (FCS) generation and checking, flow control, and reconciliation sublayer functions as required by the IEEE802.3ba specification.
Gearbox Logic

The 40G/100G Ethernet data is 64B/66B encoded in the egress direction and 66B/64B decoded in the ingress direction. Each VL from or into the PCS block is 66 bits wide. When configured at 10.3125 Gbps for 40G/100G Ethernet, the Stratix IV GT transceiver channels support a parallel data width of 40 bits.

In the egress direction, the gearbox logic is responsible for translating each 66-bit wide VL from the PCS into 40-bit wide data to the Stratix IV GT transmitter channel. In the ingress direction, the gearbox logic is responsible for translating the 40-bit wide data from each Stratix IV GT receiver channel into 66-bit wide data to the PCS.

Stratix IV GT Transceiver Configuration for 40G/100G Ethernet Implementation

The following sections discuss the required Stratix IV GT transceiver configuration and the recommended clocking for 40G/100G Ethernet implementation:

- “Required Transceiver Channel Placement”
- “Transceiver Channel Datapath Configuration” on page 8
- “40G Ethernet Transceiver Channel Datapath Clocking” on page 10
- “100G Ethernet Transceiver Channel Datapath Clocking” on page 12

Required Transceiver Channel Placement

Stratix IV GT EP4S40G2 and EP4S40G5 devices support up to six channels on each side of the device. Each channel can operate at 10.3125 Gbps. Altera recommends using these devices to implement a 40G Ethernet link.
Figure 4 shows the transceiver channel locations in EP4S40G2F40 and EP4S40G5H40 devices that support 10.3125 Gbps data rate to implement a 40G Ethernet link.

Figure 4. Transceiver Channel Locations for 40G Ethernet Implementation

Stratix IV GT EP4S100G2 and EP4S100G5 devices support up to 12 channels on each side of the device. Each channel can operate at 10.3125 Gbps. Altera recommends using these devices to implement a 100G Ethernet link.
Figure 5 shows the transceiver channel locations in EP4S100G2F40 and EP4S100G5H40 devices that support 10.3125 Gbps data rate to implement a 100G Ethernet link. Figure 6 shows the transceiver channel locations in the EP4S100G3F45/EP4S100G4F45/EP4S100G5F45 devices that support 10.3125 Gbps data rate to implement a 100G Ethernet link.

**Figure 5. Transceiver Channel Locations for 100G Ethernet Implementation in EP4S100G2F40 and EP4S100G5H40 Devices**
Figure 6. Transceiver Channel Locations for 100G Ethernet Implementation in an EP4S100G5F45 Device

Note to Figure 6:
(1) EP4S100G3F45 and EP4S100G4F45 devices only support up to 12 10G channels on each side of the device.
Transceiver Channel Datapath Configuration

In a 40G or 100G Ethernet link, each of the 4 or 10 transceiver channels operates at 10.3125 Gbps. The Stratix IV GT transceiver channel must be configured in Basic Double-Width mode with a 40-bit FPGA fabric-transceiver interface with a low latency PCS datapath, as shown in Figure 7.

Figure 7. Low Latency PCS Transceiver Configuration Required for Implementing 40G/100G Ethernet Link

For more information about the low latency PCS datapath, refer to “Low Latency PCS Datapath” in the Stratix IV Transceiver Architecture chapter in volume 2 of the Stratix IV Handbook.
Figure 8 shows the ALTGX MegaWizard™ Plug-In Manager setup for 40G/100G Ethernet implementation.

Figure 8. ALTGX Instance for 40G/100G Ethernet Implementation

- 40G Ethernet = 4 channels
- 100G Ethernet = 10 channels
40G Ethernet Transceiver Channel Datapath Clocking

Figure 9 shows the recommended transceiver channel datapath clocking for implementing a 40G Ethernet link in Stratix IV GT devices.
Transmitter Channel Datapath Clocking

The CMU0 PLL located in the transceiver block containing four channels is used to generate the high-speed serial and low-speed parallel transmitter clocks. Table 1 lists the allowed input reference clock frequencies and recommended input reference clock sources for optimum jitter performance.

Table 1. Input Reference Clock Recommendation for 40G Ethernet Link

<table>
<thead>
<tr>
<th>Input Reference Clock Frequency</th>
<th>Recommended Input Reference Clock Source for Optimum Output Jitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>644.53125 MHz</td>
<td>Dedicated refclk0 or refclk1 pin of the transceiver block containing the four 10.3125G channels</td>
</tr>
<tr>
<td>322.65625 MHz</td>
<td></td>
</tr>
</tbody>
</table>

The high-speed serial clock at 5156.25 MHz and the low-speed parallel clocks at 515.625 MHz generated by the CMU0 PLL are distributed to the four transmitter channels through the ×1 clock line. The PMA and PCS blocks within all four channels use the high-speed serial and low-speed parallel clock for their operation.

For more details about the ×1 clock network, refer to the Stratix IV Transceiver Clocking chapter in volume 2 of the Stratix IV Handbook.

The low-speed parallel clock of 515.625 MHz to each transmitter channel gets further divided-by-two at the byte serializer. This divided-by-two clock at 257.8125 MHz is forwarded as the FPGA fabric-transceiver interface clock on each channel’s tx_clkout port. You can use the tx_clkout port from one of the four channels to clock the 66-bit to 40-bit gearbox logic, as well as the PCS layer in the FPGA. Refer to “Gearbox Logic” on page 4 for more details.

Receiver Channel Datapath Clocking

As shown in Figure 9, the clock and data recovery (CDR) in each of the four receiver channels forwards the high-speed serial (5156.25 MHz) and low-speed parallel (515.625 MHz) recovered clocks to the receiver PMA and PCS blocks for their operation. The low-speed parallel recovered clock of 515.625 MHz gets further divided-by-two at the byte deserializer. This divided-by-two clock at 257.8125 MHz is forwarded as the FPGA fabric-transceiver interface clock on each channel’s rx_clkout port. You can use the rx_clkout port from one of the four channels to clock the 40-bit to 66-bit gearbox logic, as well as the PCS layer in the FPGA. Refer to “Gearbox Logic” on page 4 for more details.
100G Ethernet Transceiver Channel Datapath Clocking

Figure 10 shows the recommended transceiver channel datapath clocking for implementing a 100G Ethernet link in Stratix IV GT devices.
Transmitter Channel Datapath Clocking

The 10 channels that form the 100G Ethernet link are located across three transceiver blocks. The CMU0 PLL in each transceiver block is used to generate the high-speed serial and low-speed parallel transmitter clocks for channels located in that transceiver block.

Table 2 lists the allowed input reference clock frequencies and recommended input reference clock sources for optimum jitter performance.

For more details about the ×1 clock network, refer to the Stratix IV Transceiver Clocking chapter in volume 2 of the Stratix IV Handbook.

The low-speed parallel clock of 515.625 MHz to each transmitter channel gets further divided-by-two at the byte serializer. This divided-by-two clock at 257.8125 MHz is forwarded as the FPGA fabric-transceiver interface clock on each channel’s tx_clkout port. You can use the tx_clkout port from one of the four channels to clock the 66-bit to 40-bit gearbox logic, as well as the PCS layer in the FPGA. Refer to “Gearbox Logic” on page 4 for more details.

Receiver Channel Datapath Clocking

As shown in Figure 10, the CDR in each of the 10 receiver channels forwards the high-speed serial (5156.25 MHz) and low-speed parallel (515.625 MHz) recovered clocks to the receiver PMA and PCS blocks for their operation. The low-speed parallel recovered clock of 515.625 MHz gets further divided-by-two at the byte deserializer. This divided-by-two clock at 257.8125 MHz is forwarded as the FPGA fabric-transceiver interface clock on each channel’s rx_clkout port. You can use the rx_clkout port from one of the 10 channels to clock the 40-bit to 66-bit gearbox logic, as well as the PCS layer in the FPGA. Refer to “Gearbox Logic” on page 4 for more details.

Gearbox Requirements in 40G/100G Ethernet Implementation

The 40G/100G Ethernet link uses 64-bit/66-bit encoding to achieve DC balance on the line. The Stratix IV GT FPGA fabric-transceiver interface is 40 bits wide, as shown in Figure 9 and Figure 10. Gearbox logic is required to compensate for the data-width mismatch between the PCS (66-bit wide) and the Stratix IV GT transceiver (40-bit side).
Gearbox Requirements in the Egress Direction

In the egress direction, the MAC forwards a continuous stream of 64-bit data blocks to the PCS in the FPGA fabric. The PCS encodes the 64-bit data blocks from the MAC into 66-bit data blocks with a 64-bit/66-bit encoder, scrambles the 66-bit data blocks, and forwards a continuous stream of scrambled 66-bit data blocks to the striping sublayer within the PCS.

In a 40G Ethernet link, the striping sublayer stripes the continuous stream of 66-bit data from the PCS across four VLs. The four 66-bit wide VLs from the striping sublayer within the PCS must be mapped to four 40-bit wide physical lanes into the transmitter channels. Gearbox logic must be implemented in the FPGA fabric to do this mapping.

Figure 11 shows the recommended gearbox scheme in the egress direction for 40G Ethernet implementation.

---

Figure 11. Recommended Gearbox Scheme in Egress Direction for 40G Ethernet Implementation

In a 100G Ethernet link, the striping sublayer stripes the continuous stream of 66-bit data from the PCS across 20 VLs. The twenty 66-bit wide VLs from the striping sublayer must be mapped to ten 40-bit wide physical lanes into the transmitter channels. Gearbox logic must be implemented in the FPGA fabric to do this mapping.
Figure 12 shows the recommended gearbox scheme in the egress direction for 100G Ethernet implementation.

Figure 12. Recommended Gearbox Scheme in Egress Direction for 100G Ethernet Implementation

As shown in Figure 11 and Figure 12, each 66-bit wide VL from the striping sublayer within the PCS is clocked by the 156.25 MHz system clock. The recommended gearbox scheme implements a FIFO for each VL that decouples the system clock at 156.25 MHz and the FPGA fabric-transceiver interface clock at 257.8125 MHz. The 66-bit wide data from the PCS is written into the FIFO at 156.25 MHz and read out of the FIFO at 257.8125 MHz. Because the read clock to the FIFO is faster than the write clock, the FIFO empties if data is continuously read from it. To avoid an empty FIFO, the FIFO logic must periodically stall data being read from the FIFO. The FIFO logic must also send a data_valid signal to the gearbox logic to indicate a valid or invalid data cycle.

There is no impact on the link throughput due to stalling of FIFO reads. The stalling is only required to maintain a 10.3125 Gbps throughput on each lane of the 40G/100G Ethernet link while compensating for the frequency difference between the system clock at 156.25 MHz and FPGA fabric-transceiver interface clock at 257.8125 MHz.

In a 40G Ethernet link, the gearbox logic converts each of the four 66-bit lanes from the FIFO into four 40-bit lanes to the Stratix IV GT transmitter channels. You can use the tx_clkout (FPGA fabric-transceiver interface) signal from any one of the four channels to clock the gearbox logic and the write side of the FIFO, as shown in Figure 11.
In a 100G Ethernet link, the gearbox logic converts each of the twenty 66-bit lanes from the FIFO into twenty 20-bit lanes to the Stratix IV GT transmitter channels. The multiplexer block converts these twenty 20-bit lanes from the gearbox logic into ten 40-bit lanes to the Stratix IV GT transmitter channels. You can use the \texttt{tx_clkout} (FPGA fabric-transceiver interface) signal from any of the 10 channels to clock the multiplexer block, the gearbox logic, and the write side of the FIFO, as shown in Figure 12.

The egress-side gearbox schemes shown in Figure 11 and Figure 12 assume that the 156.25 MHz system clock and the 257.8125 MHz FPGA fabric-transceiver interface clock are both derived from the same base frequency. This can be achieved by deriving the 156.25 MHz system clock from the input reference clock with either an on-chip general purpose PLL or an off-chip clock synthesizer.

**Gearbox Requirements in the Ingress Direction**

In the ingress direction of a 40G Ethernet link, each of the four Stratix IV GT receiver channels forwards a 40-bit wide data bus to the Ethernet soft IP in the FPGA fabric. Data from each channel within the link goes to the PCS for destriping. Because each VL in the PCS is 66 bits wide, gearbox logic is required to convert the 40-bit data from each of the four channels to 66-bit data in the PCS.

Figure 13 shows the recommended gearbox scheme in the ingress direction for 40G Ethernet implementation.

**Figure 13. Recommended Gearbox Scheme in Ingress Direction for 40G Ethernet Implementation**

In the ingress direction of a 100G Ethernet link, each of the 10 Stratix IV GT receiver channels forwards a 40-bit wide data bus to the Ethernet soft IP in the FPGA fabric. Data from each channel within the link goes to the PCS for destriping. Because each of the 20 VLs in the PCS is 66 bits wide, gearbox logic is required to convert the 40-bit data from each of the 10 channels to 66-bit data in the PCS.
Figure 14 shows the recommended gearbox scheme in the ingress direction for 100G Ethernet implementation.

Figure 14. Recommended Gearbox Scheme in Ingress Direction for 100G Ethernet Implementation

As shown in Figure 13 and Figure 14, each receiver channel in a 40G or 100G Ethernet link forwards a 40-bit wide data to the FPGA fabric.

Altera recommends enabling the \texttt{rx\_coreclk} port (read clock for the receiver phase compensation FIFO) in the ALTGX MegaWizard Plug-In Manager instance used to instantiate the 40G/100G Ethernet link. Drive each channel's \texttt{rx\_coreclk} port with the recovered clock \texttt{rx\_clkout} from any one of the 4/10 receiver channels in the 40G/100G Ethernet link. Use this \texttt{rx\_clkout} signal to clock the gearbox logic for all receiver channels, as shown in Figure 13 and Figure 14.

For more information about using \texttt{rx\_coreclk} ports, refer to “User-Selected Receiver Phase Compensation FIFO Read Clock” in the \textit{Stratix IV Transceiver Clocking} chapter in volume 2 of the \textit{Stratix IV Handbook}.

Because the gearbox logic takes in 40-bit data and reads out 66-bit data at 257.8125 MHz, it could become empty. To avoid this, the gearbox logic must periodically stall reading out data. To indicate this periodic stalling of data to the FIFO, the gearbox logic sends a \texttt{data\_valid} signal to the downstream FIFO. The FIFO must register the 66-bit data only when the \texttt{data\_valid} signal from the gearbox logic is asserted. The FIFO reads out the 66-bit data to the PCS with the 156.25 MHz system clock, as shown in Figure 13 and Figure 14.
The recovered clock (rx_clkout[0]) at 257.8125 MHz and the system clock at 156.25 MHz can have a maximum of ±100 PPM difference. You must design the FIFOs shown in Figure 13 and Figure 14 to be able to compensate for this PPM difference between the two clock domains.

Summary

Stratix IV GT devices offer a single-chip solution for implementing 40G/100G Ethernet MAC devices in a 40G/100G line card. This application note provides guidance on the implementation of 40G/100G Ethernet links in a Stratix IV GT device.

Document Revision History

Table 3 shows the revision history for this document.

<table>
<thead>
<tr>
<th>Date</th>
<th>Version</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 2011</td>
<td>1.2</td>
<td>- Updated reference to Altera’s IP partners.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Updated document layout.</td>
</tr>
<tr>
<td>December 2009</td>
<td>1.1</td>
<td>- Updated figures to remove 8G channels.</td>
</tr>
<tr>
<td>June 2009</td>
<td>1.0</td>
<td>- Initial release.</td>
</tr>
</tbody>
</table>