Hardware Acceleration in SoC FPGAs

Introduction

One of the key benefits of integrating a processor and FPGA into a single device is the ability to accelerate system performance by offloading critical functions to the FPGA. Transferring the data quickly and coherently is key to realizing this performance boost. The integration of an ARM processor and FPGA logic with high speed, on-chip interconnect buses for performance, along with an Accelerator Coherency Port for coherency, makes this possible in the SoC FPGA-based systems of today.

This Architecture Brief describes the merits of Altera SoC FPGAs' inclusion of an ARM Cortex-A9 processor, and a highly-versatile Accelerator Coherency Port, to accelerate operations in a wide range of applications.

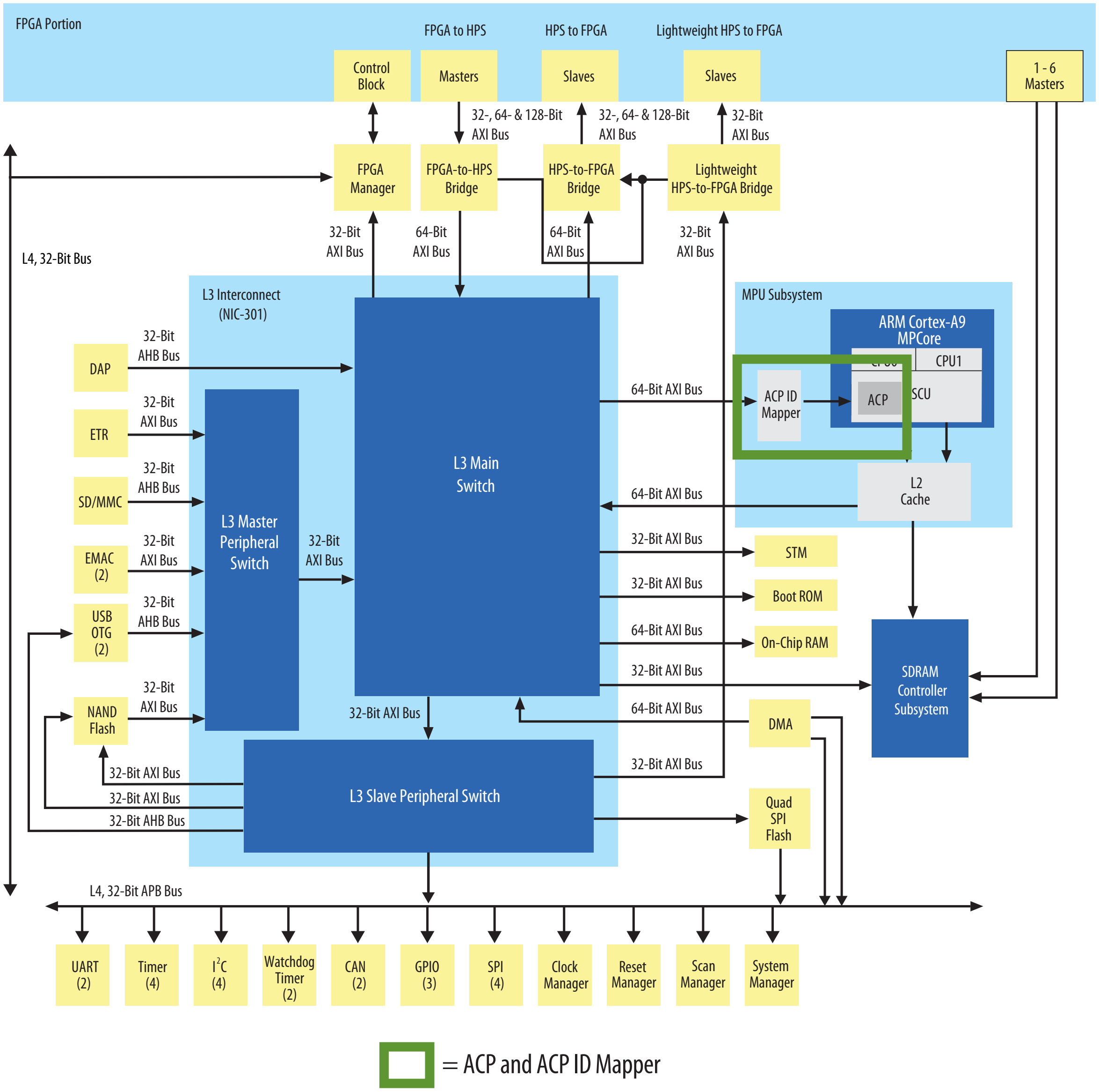
Key aspects of this Architecture Brief are highlighted in an online video: "Processor to FPGA Interconnect", which can be found at www.altera.com/socarchitecture.

Hardware Acceleration and Cache Coherency

A key potential benefit of the integrated processor and FPGA system is the ability to boost system performance by accelerating compute-intensive functions in FPGA logic. The processor can be offloaded by accelerating practically anything in FPGA logic—from calculating a cyclic-redundancy check (CRC) to offloading the entire TCP/IP stack. When the FPGA-based accelerator produces a new result, the data needs to be passed back to the processor as quickly as possible, so that the processor can update its view of the data.

ARM Cortex-A9 processor-based SoC FPGAs include a feature called an Accelerator Coherency Port (ACP). Through the ACP, new data produced by an FPGA-based hardware accelerator is transferred directly to the processor's L2 cache, via a low-latency direct connection ***(Figure 1)***. This operation is performed not just quickly, but coherently too.

Figure 1 Altera Cyclone V SoC FPGA block diagram with the Accelerator Coherency Port (ACP) and ACP ID Mapper highlighted



The ACP logic automatically maintains L2 cache coherency, so a coherent data transfer requires approximately 30 cycles. The alternative method to ensure data coherency is to flush the L2 cache, which requires hundreds of cycles to complete. Altera SoC FPGAs support coherent transactions for both FPGA-based functions and for processor peripherals, as shown in ***Table 1***. Other SoC FPGAs only support FPGA functions via a single dedicated port and do not support transactions from processor peripherals.

ARM originally designed the ACP interface for full-custom SoCs, which generally have only a few dedicated accelerators or a few peripherals that require ACP support. Consequently, the ARM ACP interface only supports eight transactions, in flight or pending. However, because of the SoC FPGA's flexible and programmable architecture, there may be many more hardware accelerators that require coherent support. To support more than eight functions, Altera SoC FPGAs incorporate an ACP ID mapper that supports an unlimited number of pending transactions with any eight transactions currently in flight.

Table 1: Accelerator Coherency Port Differences in SoC FPGAs

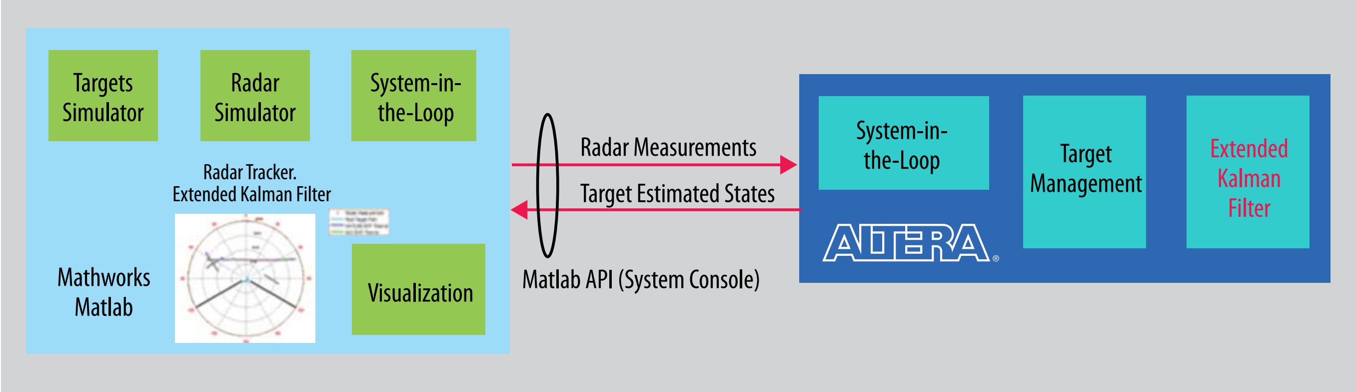
	Altera SoC FPGAs	**Xilinx Zynq-7000 EPP**
FPGA-Based Masters Supported by ACP	Yes	Yes
Processor Peripheral Masters Supported by ACP	Yes	No
ACP ID Mapper	Yes	No
ACP In-Flight Transactions Supported	8	8 total in flight or pending
ACP Pending Transactions Supported	Unlimited	8 total in flight or pending
ACP Port Configuration	x64 AXI	x64 AXI
ACP Port Clock Source	½ CPU Clock (400 MHz for 800 MHz CPU)	FPGA (150 MHz)

Application Example: Extended Kalman Filter

The Altera Extended Kalman Filter (EKF) reference design provides an example of the benefits of implementing hardware acceleration in the FPGA. The EKF is an algorithm commonly used in military radar, sonar, guidance and navigation systems, and inertial navigation sensors; as well as automotive sensor fusion and industrial motor control. The EKF is the non-linear version of the Kalman Filter that is suited to work with systems whose model contains non-linear behavior. The algorithm linearizes the non-linear model at the current estimated point in an iterative manner as a process evolves.

Hardware acceleration of this algorithm can be realized by offloading the generic portions of the algorithm to the FPGA while retaining the application specific portions on the ARM processor. This approach can provide a >2x system performance improvement while utilizing <10% of the FPGA logic in a 110KLE Cyclone V SoC FPGA.

Figure 2. Altera Extended Kalman Filter Reference Design Block Diagram - Radar Implementation



Conclusion

Offloading key functions from the processor to the FPGA can result in substantial system performance increases as well as decreasing system power. High speed / low latency interconnect between the processors and FPGA is critical for achieving this performance benefit in SoC FPGA-based systems; for accelerator functions which utilize the L2 cache, it is equally critical to maintain memory coherency through the use of an Accelerator Coherency Port (ACP). Not all ACPs in SoC FPGAs on the market today are created equal. The ACP in Altera SoCs implements a wider range of options and masters for maximum versatility and performance.

Want to Dig Deeper?

The Extended Kalman Filter reference design can be found on Altera's website at:
http://www.altera.com/literature/ds/Extended_Kalman_Filter.pdf
More information on development tools for implementing hardware acceleration on Altera SoC FPGAs such as the ARM DS-5 Streamline profiling tool, which can be used for identifying hotspots in the code as candidates for FPGA offload, and the Altera SDK for OpenCL™, used to abstract the FPGA and implement parallelization, can be found here:

http://ds.arm.com/ds-5/optimize/

http://www.altera.com/products/software/opencl/opencl-index.html

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

Altera European Headquarters
Holmers Farm Way
High Wycombe
Buckinghamshire
HP12 4XF
United Kingdom
Telephone: (44) 1494 602000

Altera Japan Ltd.
Shinjuku i-Land Tower 32F
6-5-1, Nishi-Shinjuku
Shinjuku-ku, Tokyo 163-1332
Japan
Telephone: (81) 3 3340 9480
www.altera.co.jp

Altera International Ltd.
Unit 11- 18, 9/F
Millennium City 1, Tower 1
388 Kwun Tong Road
Kwun Tong
Kowloon, Hong Kong
Telephone: (852) 2 945 7000
www.altera.com.cn