With high-speed 100-GbE communication network standards converging, switching functions play a key role in the smooth functioning of the Internet. The aggregated network traffic doubles every six months and grows in complexity as it is transported across multiple protocols, challenging the limits of current switch architectures. Because today's single-chip-based architectures are unable to meet this demand for increased bandwidth and complexity, there is a need to develop efficient algorithms and switching architectures to meet the high-speed network requirements. Altera® Stratix® V FPGAs enable hardware designers to integrate true 100-GbE components for next-generation switches and routers that ensure QoS while balancing the distribution of data through the system.

Introduction

Switching networks are used as intelligent interconnect structures for computer networks and high-speed 100-GbE networks. Historically, these interconnect networks were deployed independently in telephone, computer, and internet infrastructure networks.

Today's communication networks are converging towards a common interconnect switching technology with the objective of a single interconnect network capable of distributed processing capability. Traditional telephony networks, based on time-division multiplexed (TDM) switching, are transitioning to Ethernet/IP packet-based switching networks. This transition has an impact on the switching function as Ethernet frames and IP packets have variable packet lengths that add further complexity to the buffering and segmentation of the switch. This evolution towards IP switching requires the support of existing networks and demands hybrid switches.

Hybrid switch fabric devices support different protocols, such as Ethernet, SONET, and TDM, and are even used for backplane switching. These devices must also support various interconnect protocols, such as SPI-4/-5.2, Interlaken, SRIO and XAUI, in addition to different encoding schemes, such as 8B/10B and 64B/66B. Inside the switch, these packets are stored in buffers and later segmented into different packet sizes based on the switch algorithm to obtain better overall network traffic efficiency.

Figure 1 shows a typical switch chassis consisting of a number of switch line cards connected to network processing units (NPUs) that are interconnected with a switch fabric. The data enters through an ingress processor, goes through the switch fabric, and leaves from the egress processor. The primary function of the NPUs is to inspect the arriving packets to determine their quality of service (QoS), perform any packet modification such as encapsulation—which is protocol dependent—and encryption/decryption to ensure network security.

July 2010   Altera Corporation

Feedback   Subscribe

Figure 1. Typical Switch Chassis
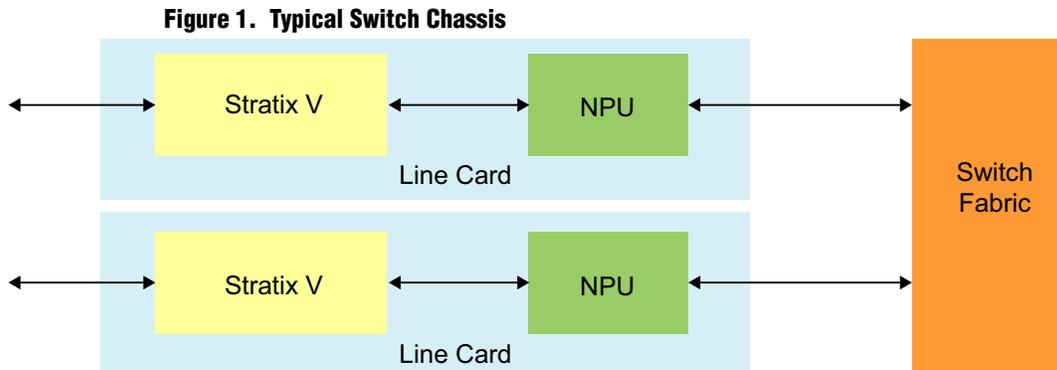
This switching architecture has an inherent memory limitation that does not meet the requirement for the packet queuing process. Even in the shared memory architecture, where switches contain a global memory to be shared across all the line cards for queuing purposes, the performance is directly scaled to the memory as the demand for bandwidth grows.

Because compensating for the memory challenges and redesign with a new line card using ASIC or ASSP platforms is an expensive proposition, designers are always looking for new cost-effective alternatives for their switch fabric architectures. One way to overcome this challenge is to queue the packets at the input stream, a method known as the input-queuing model. In this architecture, the fabric only provides a transport between the ingress and egress NPUs and acts like a typical crossbar switch. FPGAs, such as Altera's Stratix V devices, are well suited for this implementation and provide a cost-effective and a faster-time-to-market solution.

Crossbar Switching

A basic crossbar or crosspoint switch, shown in Figure 2, has a set of inputs, a corresponding set of outputs, and a set of addresses mapping inputs to outputs.

Figure 2. Crossbar-Switching Example
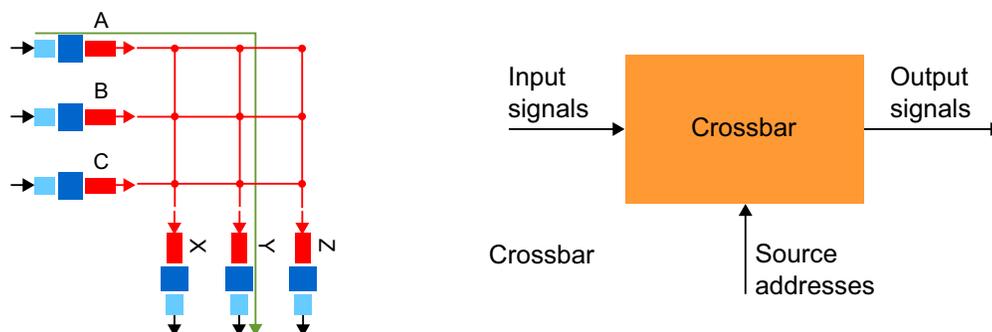
Crossbars can be classified as:

■ Blocking

■ Non-blocking

Blocking occurs when the progress of one message through the crossbar switch is stalled by another message that is not destined for the same output port. Figure 3 shows the blocking architecture of the crossbar switch, where the connection from I1 to O2 inhibits the path from I2 to O1.

Figure 3. Blocking Architecture

I1 → Active → O1
I2 → Blocked → O2
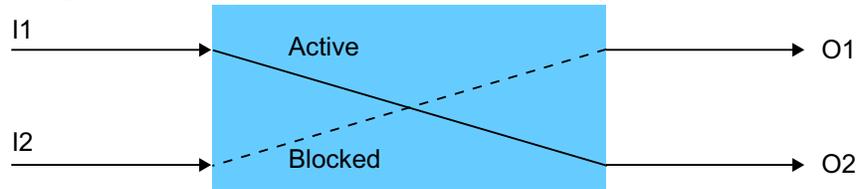
Non-blocking occurs when the connection path is independent of any other existing connections. Non-blocking switches are not scalable but provide higher performance. In contrast, blocking switches ease the port buffering challenges within the switch architecture. Figure 4 shows the non-blocking architecture of the crossbar switch, where traffic from I1 to O2, as well as I2 to O1, can take place simultaneously.

Figure 4. Non-Blocking Architecture

I1 → Active → O1
I2 → Active → O2
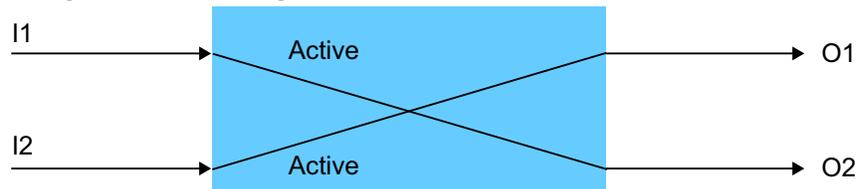
A FPGA platform provides flexibility to the system designer to implement both the blocking as well as non-blocking architectures on a single switch device. A crossbar switch implemented on a FPGA platform is able to support data transported from a wide range of network types, including variable-length IP packets, ATM, and SONET streams. Irrespective of the packet type being transported, the QoS is not compromised and can be customized based on the design or packet type requirement.

As shown in Figure 5, an external arbitrator logic is used to determine and control the QoS based on the input queue status. The arbiter chooses a connection within the crossbar switch and determines the order and priority in which packets are routed through the network. Due to their high I/O port count, Stratix V FPGAs are well suited for this implementation of a memoryless crossbar switch. Designers are given the flexibility to implement various arbitration algorithms to manage the QoS and maximize the efficiency of the fabric.

Figure 5. Crossbar Switch with Arbitration

Input queue buffer → Crossbar → Output queue buffer
Arbiter logic
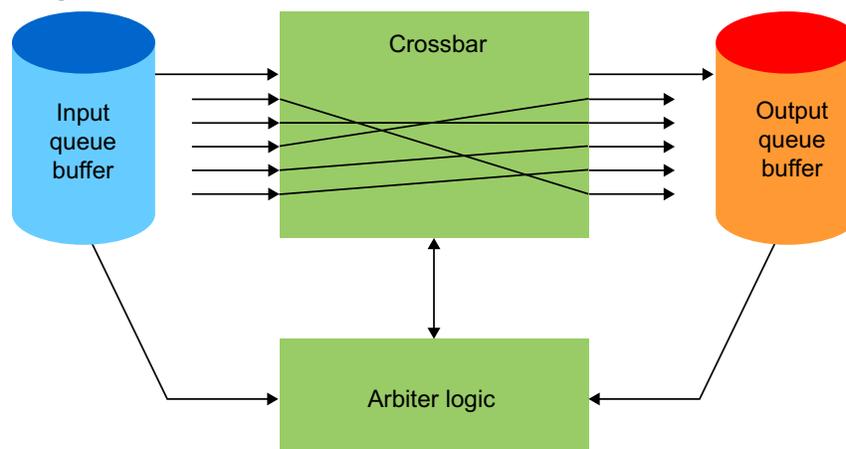
Backplane Switching

A backplane switch fabric (Figure 6) is a specialized form of switch fabric dedicated to a single task—connecting multiple line cards via a switch architecture across a system backplane. Data traffic is routed from one card to another through the switch. A switch can have multiple backplanes and even multiple cards in a chassis. Line cards and a central switch function make up a complete switch-router system.

Figure 6. Backplane Switch
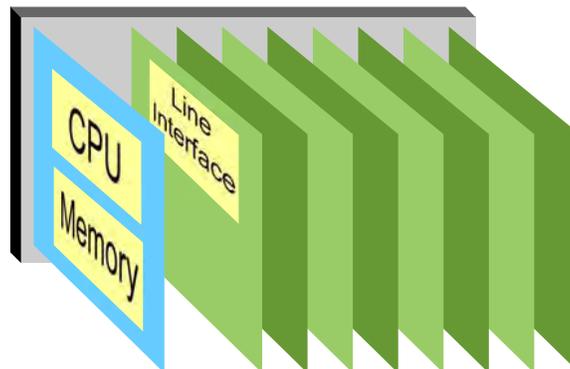
The line cards, which typically manage the PHY/MAC functions, take in line-speed data and send it to the packet processing module of the NPU or traffic manager. The output traffic is then sent to a switch line card to go through the switching function, which consists of a queue manager and a switching mechanism to route the data path. The switch line card connects the other line cards within a chassis via the backplane infrastructure.

Backplane switch fabrics provide the switching between arrays of line cards for network switches and routers. Deployed as chipsets, they provide cost-effective performance for Layer 2 and Layer 3 applications. However, these chipsets are not flexible enough to allow dynamic reconfiguration of the chassis based on the client requirements. These chipsets are also unable to handle changes to the protocol line rates. In most cases, the line card must be physically replaced if there are protocol changes to the client or overall data patterns.

Stratix V FPGAs with 66 full-duplex transceiver ports provide an ideal switching platform built on a lossless and memoryless switch. With the partial reconfiguration feature, operators can dynamically reconfigure the line card or backplane switch to meet client requirements. This configuration can be done remotely without affecting the operation of other clients on the same chassis. When a packet transfer is granted, the switch connects the packet to the proper output port.

The switch implemented on Stratix V FPGAs performs limited queuing functions using its 50-Mb embedded memory. Most of the queuing function is performed by an external queue manager. The FPGA acts as a simple crossbar switch with high port aggregation and backplane connectivity. The transferring and queuing of packets are implemented before the actual switching function, and all output buffering is done at the back end of the queue manager.

Implementing Crossbar Switching and Backplane Switching

Altera's 28-nm Stratix V FPGAs are designed with a number of significant innovations that lead to the highest bandwidth, highest system-level integration, and greatest flexibility available today in a single FPGA. The implementation of crossbar-switching or backplane-switching functions makes use of the following new and enhanced device features:

- 66 integrated serial transceivers covering a continuous range from 600 Mbps to 12.5 Gbps, supporting both line sides, client-side backplane interfaces, and next-generation serial knowledge base processors (KBPs)

- 1.6 Tb of total serial switching capability provides the necessary high port count required for 400-GbE system designs

- Stratix V FPGAs offer up to 460 Gbps of memory bandwidth supporting six 72-bit DDR3 memory interfaces at 800 MHz, along with up to 50 Mb of embedded memory on the device

- High-performance, high-density, low-power FPGA fabric with redesigned adaptive logic modules (ALMs) and enhanced routing structures, providing the throughput necessary to support wide parallel buses in the user logic

- Partial reconfiguration of core logic, allowing individual switching client port and logic to be added, dropped, or modified, without affecting the operations of other clients or ports

- Dynamic reconfiguration of transceivers, allowing changes to the analog physical medium attachment (PMA) settings for every transceiver without affecting the operations of any other transceivers or the rest of the device

Multiple serial switching bandwidth configurations of up to 1.6 Tb can be implemented in a Stratix V FPGA device. Figure 7 shows the diagram of a crossbar switch implemented on a Stratix V FPGA. The shared memory shown in the diagram is the DDR3 interface block. An external ternary content addressable memory (TCAM)/KBP ASSP can also be used to interface the device.

**Figure 7. Crossbar Switch Using a Stratix V FPGA**
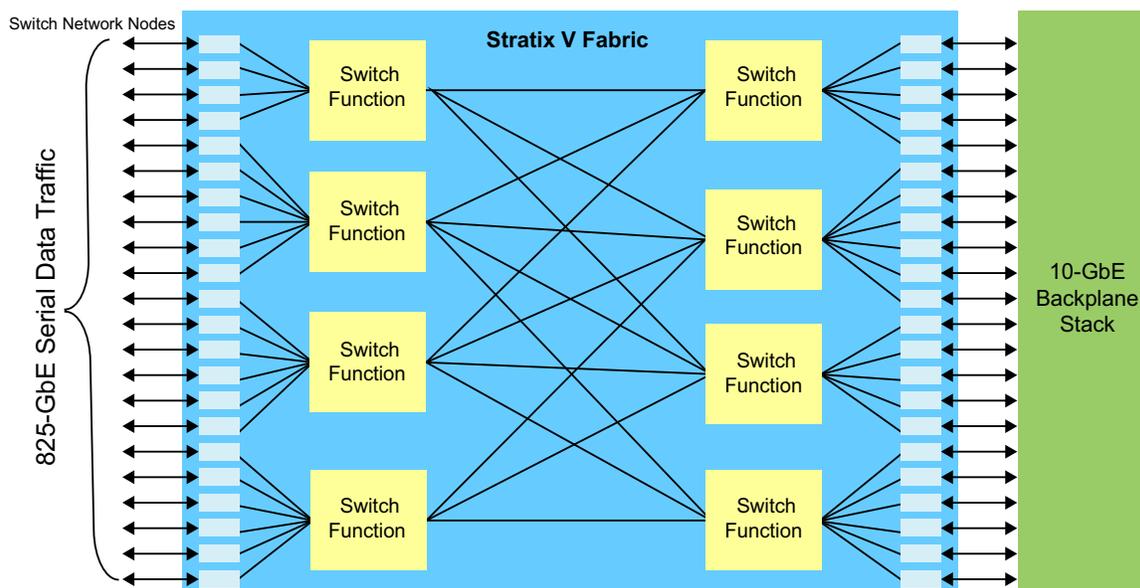
Integrated Serial Transceivers

Stratix V FPGAs have up to 66 transceivers operating at data rates from 600 Mbps to 12.5 Gbps, with another set of dedicated transceivers that support line rates from 20 Gbps to 28 Gbps. Each transceiver includes programmable pre-emphasis, equalization and differential output voltage, and a comprehensive physical coding sublayer (PCS) that includes support for a wide range of industry standard and proprietary protocols.

Figure 8 shows the diagram of a backplane switch implemented on a Stratix V FPGA. For backplane-switching applications, the line-side interface, which is connected to other line cards on the chassis, uses a group of these 66 transceivers. Stratix V FPGAs include low jitter inductor-capacitor (LC) transmit phase-locked loops (PLLs) that are ideal for line-side protocols. The line-side data rate depends on the protocol supported, usually from 600 Mbps to 12.5 Gbps.

**Figure 8. Backplane Switch Using a Stratix V FPGA**
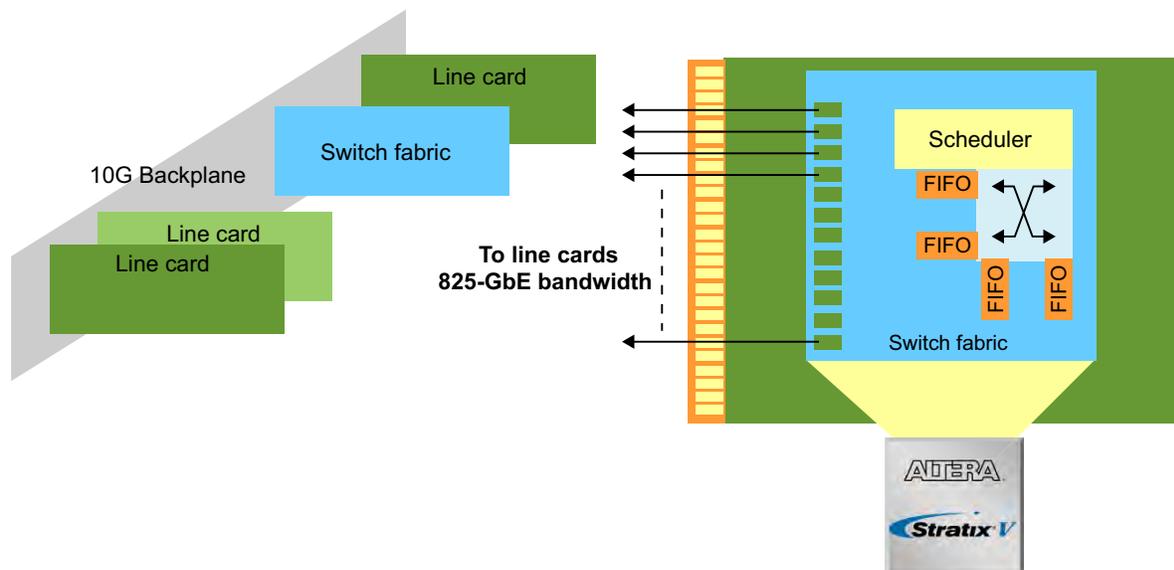
The client-side interface is normally connected to a backplane or optical module. The transceiver clocking architecture in Stratix V devices allows for the independent clocking of channels as well as the channel bonding required by the line-side interface. The client-side data rates depend on the various client protocols, but usually are from 600 Mbps to 12.5 Gbps with 10-GbE backplane support. With an operating range of up to 28 Gbps, the transceivers in Stratix V FPGAs are ideal for this application.

Partial Reconfiguration of Core Logic and Dynamic Reconfiguration of Transceivers

Because a FPGA implementation of a backplane switch in the router chassis has various line cards or clients connected to it, the bandwidth allocated to these line cards can vary depending on the quality of service, while the rest of the system continues to carry traffic and operate as usual. There is also a possibility of dynamically changing the number of line cards based on overall chassis bandwidth. All of these changes must be made without affecting overall operation of the system.

To accommodate these system changes, Stratix V FPGAs offer partial reconfiguration of the FPGA core and dynamic reconfiguration of the serial transceivers. Both techniques, shown in Figure 9, are non-intrusive. The rest of the FPGA continues to operate as usual while the changes are taking place, thus maximizing the system uptime.

Figure 9. Partial Reconfiguration of FPGA Core and Dynamic Reconfiguration of Serial Transceivers
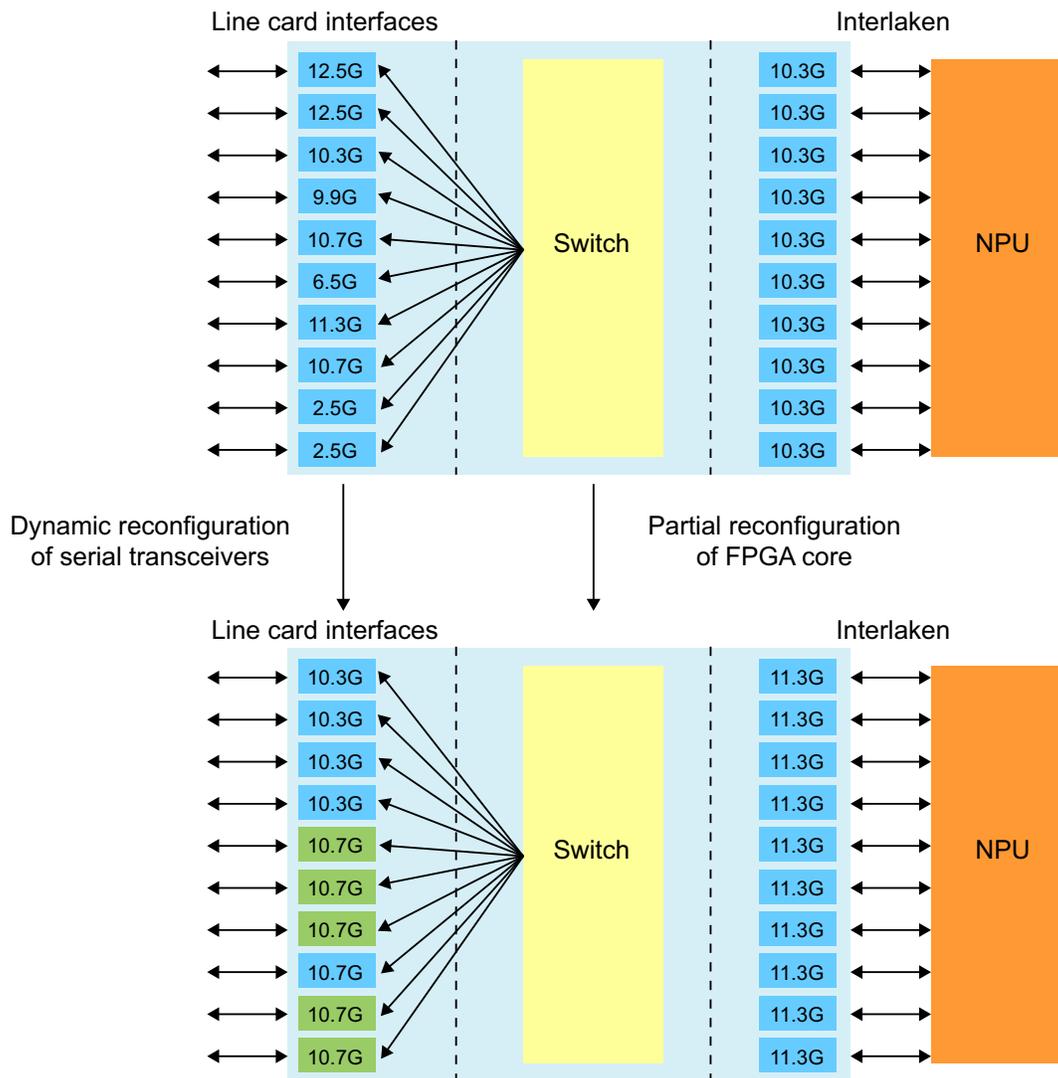
To add a line card, the transceiver is enabled and configured using dynamic reconfiguration. The protocol logic and multiplexing functions are added into the FPGA core using partial reconfiguration of the core. To remove an unused line-card interface, the transceiver is disabled and the client logic is removed from the FPGA core, thus reducing the overall power dissipation.

High-Performance, Low-Power FPGA Core Fabric

The implementation of a multi-100-GbE design in a FPGA requires a core fabric that can support pipelined, high-frequency parallel buses with low power dissipation. Stratix V FPGAs includes a number of key innovations that address these requirements:

■ Redesigned ALMs that double the number of registers available, providing easier timing closure for register-rich and heavily pipelined designs

■ An enhanced MultiTrack routing architecture with more connections to neighboring logic elements, providing increased system performance, higher logic utilization, and reduced compile times for tightly packed designs

■ New 20-Kb internal memory blocks with high performance and a small footprint, including built-in error correcting code (ECC) protection

■ Programmable Power Technology that automatically reduces the static power dissipation in the core for non-critical timing paths

■ Clock trees that can be turned on and off in sections to minimize dynamic power dissipation and prevent unnecessary switching noise

■ High-performance, high-K metal gate 28-nm process technology, delivering state-of-the-art performance and power efficiency with an 0.85-V core supply voltage

Conclusion

The explosive demand for bandwidth in data networking is driving service providers to implement new switching and routing algorithms. Device flexibility is the key in the implementation of these new algorithms. Recognizing this market need, Altera has included a number of key innovations in Stratix V FPGAs that directly address the needs of high bandwidth switching and routing applications. These innovations allow system designers to reduce cost, power, and board space while enjoying the flexibility of a FPGA-based system with partial and dynamic reconfiguration capability. In addition, Altera provides the industry-leading Quartus® II design software, reference designs, partner IP solutions, and Stratix V hardware development kits.

Further Information

- Stratix V FPGAs: Built for Bandwidth:
  www.altera.com/products/devices/stratix-fpgas/stratix-v/stxv-index.jsp

- Literature: Stratix V Devices:
  www.altera.com/products/devices/stratix-fpgas/stratix-v/literature/stv-literature.jsp

Acknowledgements

- Rishi Chugh, Senior Product Marketing Manager, Altera Corporation

Document Revision History

Table 1 shows the revision history for this document.

Table 1. Document Revision History

Date	Version	Changes
July 2010	1.1	Minor text edits.
April 2010	1.0	Initial release.