White Paper

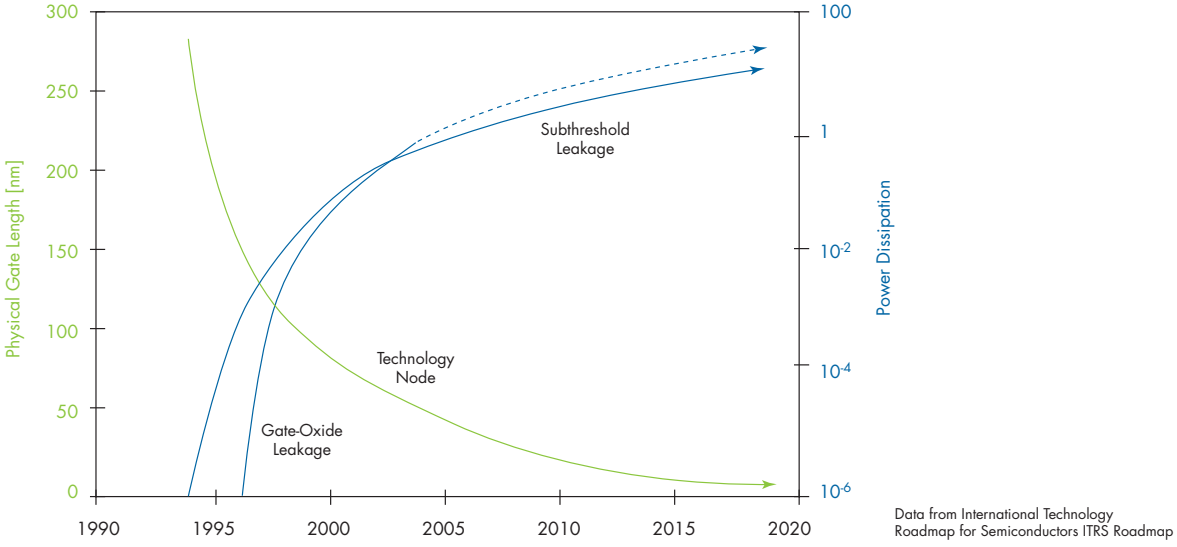
Stratix III Programmable Power

Introduction

Traditionally, digital logic has not consumed significant static power, but this has changed with very small process nodes. Leakage current in digital logic is now the primary challenge for FPGAs as process geometries decrease. While the move to the 65-nm process delivers the expected Moore's law benefits of increased density and performance, the performance increases can result in significant increases in power consumption, introducing the risk of consuming unacceptable amounts of power.

If no power-reduction strategies are employed, power consumption becomes a critical issue because static power can increase dramatically with the 65-nm process. Static power consumption rises largely because of increases in various sources of leakage current. Figure 1 shows how these sources of leakage current (shown in blue) increase as the technology makes smaller gate lengths possible (shown in green). In addition, without any specific power optimization effort, dynamic power consumption can increase due to the increased logic capacity and higher switching frequencies that are attainable.

Figure 1. Static Power Dissipation Increases Significantly at Smaller Process Geometries



Power consumption is composed of static and dynamic power. Static power is the power consumed by the FPGA when it is programmed with a Programmer Object File (**.pof**) but no clocks are operating. Both digital and analog logic consume static power. In an analog system, static power is primarily composed of the quiescent current of the analog circuit based on its interface configuration. The sources of static leakage current in 65-nm are shown in Figure 2 and Table 1.

Figure 2. Transistor Leakage Diagram

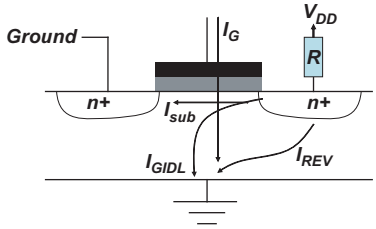
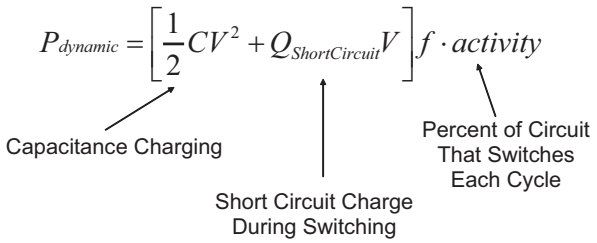


Table 1. Sources of Static Power

	Impact	Sensitivity	Design Techniques
Subthreshold (weak inversion) leakage (I_{SUB})	Dominant	Supply voltage Gate threshold voltage Temperature Channel length	Reduced core voltage Increased voltage threshold Increased gate lengths
Gate-induced drain leakage (I_{GIDL})	Small	Gate oxide thickness Supply voltage	Dual gate oxide
Gate direct-tunneling leakage (I_G)	Small	Gate oxide thickness Supply voltage	Dual gate oxide
Reverse-biased junction leakage current (I_{REV})	Negligible	N/A to low voltage CMOS	None required

Dynamic power is the additional power consumed through the operation of the device caused by signals toggling and capacitive loads charging and discharging. As shown in Figure 3, the main variables affecting dynamic power are capacitance charging, the supply voltage, and the clock frequency. Dynamic power decreases with Moore's law by taking advantage of process shrinks to reduce capacitance and voltage. The challenge is that more circuits are implemented with each process shrink and the maximum clock frequency increases. While the power reduction declines for an equivalent circuit from process node to process node, the FPGA capacity keeps doubling and the maximum clock frequency keeps increasing.

Figure 3. Variables Affecting Dynamic Power



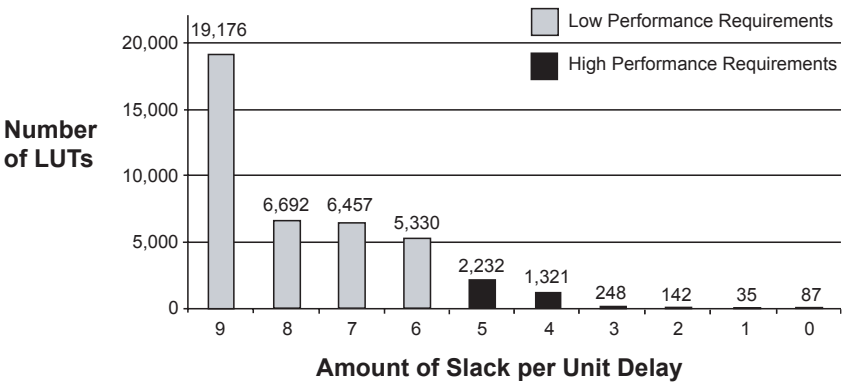
Stratix III Architecture

Altera® Stratix® III FPGAs attack these power challenges with innovative architecture, along with the latest advancements in process technology and circuit techniques.

Programmable Power Technology

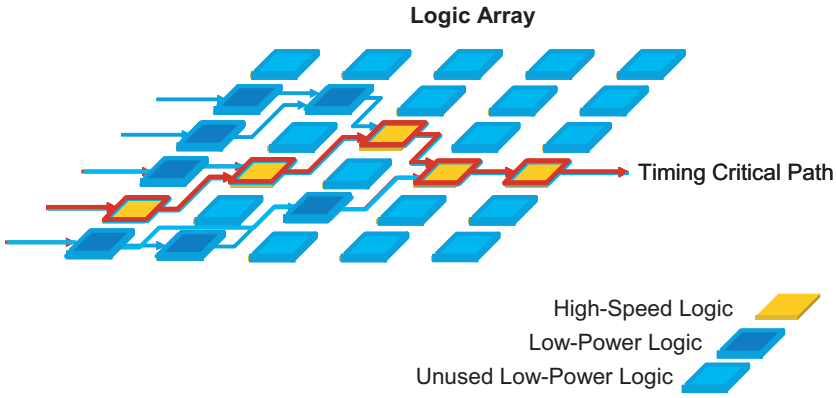
With “Programmable Power Technology,” Altera introduces a radical and unprecedented method for reducing power in high-end FPGAs. Traditionally, all high-performance FPGAs are implemented with a high-performance fabric where every logic element (LE) provides the maximum performance with a subsequent high leakage power. Altera's new Programmable Power Technology takes advantage of the fact that most circuits in a design have excess slack and therefore do not require the highest performance logic. Figure 4 shows a typical excess slack histogram where the majority of the paths (on the left) have slack and only a few critical paths (on the right) need the highest performance logic to meet timing requirements.

Figure 4. Example Slack Histogram Showing a Small Number of Circuits With Little or No Slack



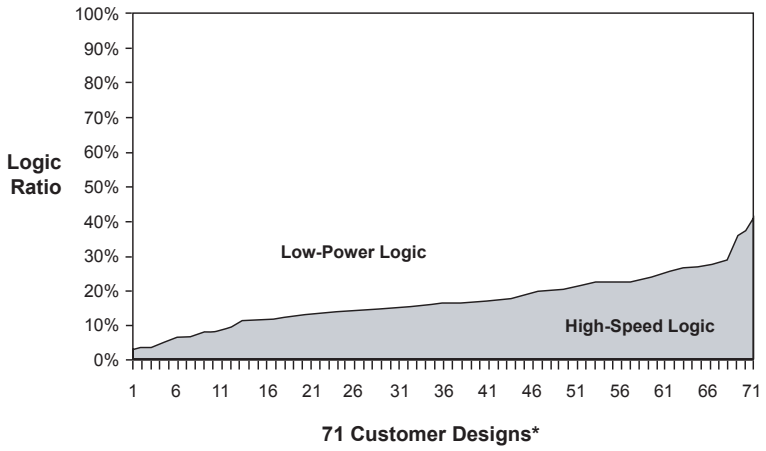
Programmable Power Technology enables the Stratix III logic fabric to be programmed at the logic array block (LAB) level to provide high-speed logic or low-power logic, depending on what is required by the specific logic path (shown in Figure 5). In this way, the small percentage of circuits that are timing critical get the high-speed setting, with the rest using the low-power setting, resulting in a dramatic 70 percent decrease in leakage power for the low-power logic. In addition, placing unused logic, as well as digital signal processing (DSP) blocks and TriMatrix memory, into low-power modes further decreases power. Stratix III Programmable Power Technology enables an optimal combination of high-speed logic to achieve the desired system performance while the rest of the logic is put into low-power mode, minimizing leakage current and resulting in the lowest power possible.

Figure 5. Stratix III Programmable Power Technology



Altera engineers analyzed benchmarks across 71 designs to evaluate the amount of high-speed logic that is typically required for a design. They compiled these designs to meet the highest performance that could be achieved within the FPGA fabric. Across these 71 designs, the average amount of high-speed logic required was about 20 percent, as shown in Figure 6.

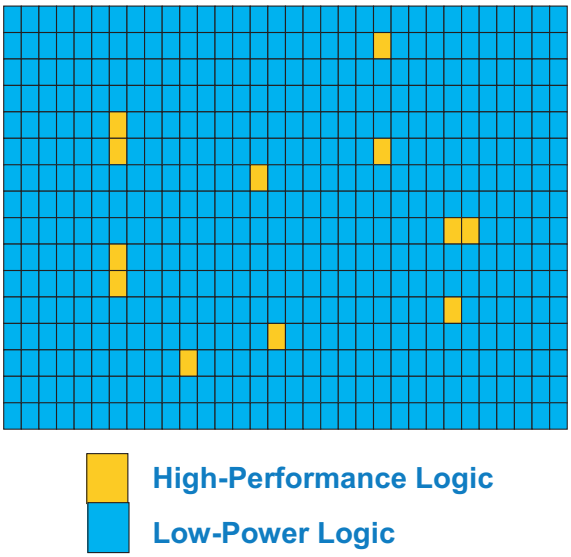
Figure 6. Benchmarks of High-Speed vs. Low-Power Logic Requirements



These benchmarks ranged from 5 to 40 percent utilization of high-speed logic when the absolute highest performance was required from the logic fabric. If more high-speed logic was applied to the designs, no more performance could be obtained because the critical paths of the designs were limited by the highest performance logic available in the FPGA. In many applications, however, Altera has found that customer designs are not performance-limited. In cases where performance requirements are 15 to 20 percent less than the highest achievable f_{MAX} in the Stratix III fabric, then most to all of the high-speed logic is replaced by low-power logic, further reducing static power.

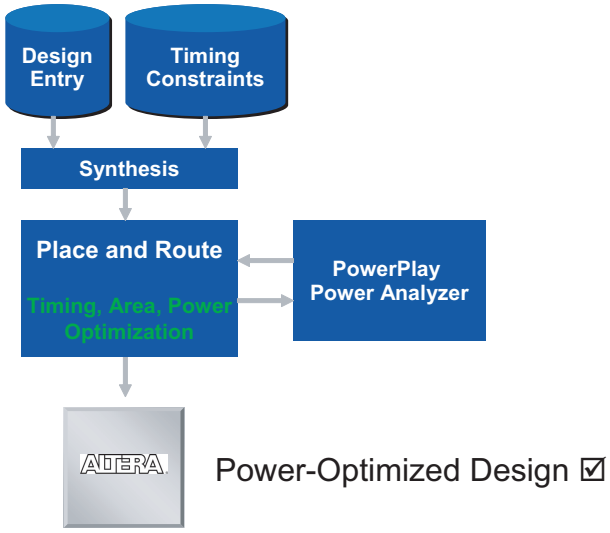
The ability to deliver the exact amount of high-speed logic required for a design to reach its desired performance can be controlled with a very high degree of precision. The programmability between high-speed and low-power logic is controlled on a per-tile basis (each tile contains two LABs, or a LAB and DSP block, or a TriMatrix memory, all with associated routing). On the largest Stratix III FPGA, over 5,000 tiles can be individually controlled as high speed or low power to get the lowest possible power for the design (see Figure 7). Altera's Quartus® II development software automatically optimizes the design by placing tiles into high-speed and low-power mode, requiring no user effort.

Figure 7. Stratix III Tile Array With Individual Programmability Between High-Speed and Low-Power Logic



Each time the Quartus II software compiles a design for a Stratix III FPGA, it automatically optimizes the design to meet specified timing constraints while minimizing power. The resulting programming file is loaded into the FPGA and includes information that sets each tile into its high-speed or low-power configuration, as shown in Figure 8. The final programming of tiles for high speed or low power is fully visible in Quartus II software.

Figure 8. Quartus II Design Flow Including Automatic Power Optimization



Selectable Core Voltage

Stratix III selectable core voltage allows the designer to use a 0.9V or 1.1V core voltage based on performance requirements. The 0.9V core voltage provides the overall minimum dynamic and leakage power, while the 1.1V core voltage delivers the overall highest performance. Dynamic power scales with the square of core voltage while static power scales by the power of 2.5 of core voltage as shown in Table 2.

Table 2. Stratix III Power Compared to Stratix II Power Across Selectable Core Voltage

Core Voltage	Dynamic Power Reduction From 1.2V	Static Power Reduction From 1.2V
1.1V	33%	52%
0.9V	55%	64%

The selectable core voltage input can be set to 0.9V or 1.1V during board design. This core voltage supplies all the LABs, memories, and DSP functions in the core fabric. The selectable core voltage affects the Stratix III fabric performance, so when a device and speed grade are selected in Quartus II software, a core voltage selection is also required. Quartus II software uses timing and power models specific to the selected core voltage to implement all timing-dependent and power-dependent analysis and optimization.

When choosing which core voltage to use, the designer must take into account the system performance requirements reported from Quartus II timing analysis. If the system performance requirements can be met with 0.9V, it will always produce lower power than with 1.1V.

Combining Programmable Power Technology and Selectable Core Voltage

In combination, the Programmable Power Technology and selectable core voltage deliver various performance and power operating points. Figure 9 shows that, even at the 1.1V core voltage setting, static power is significantly lower than for previous-generation devices. In addition, static power varies considerably depending on the utilization of the

various resources, such as DSP blocks and TriMatrix memory blocks. Figure 9 shows three lines for Stratix III static power: maximum leakage, medium leakage, and low leakage, which are further defined in Table 3.

Figure 9. Stratix III Static Power vs. Utilization and Performance]

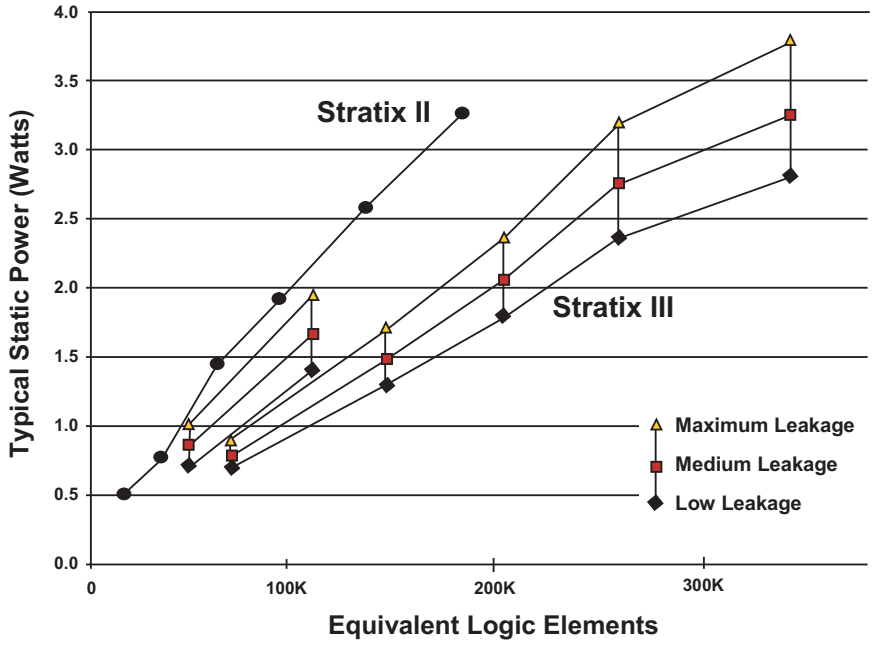


Table 3. Stratix III Operating Conditions Shown in Figure 9

	% LAB Utilization	% High-Speed Mode	% RAM/DSP Used
Maximum Leakage	100%	25%	100%
Medium Leakage	100%	10%	50%
Low Leakage	100%	0%	0%

The combined static and dynamic power varies across combinations of core voltage and percentage of high-speed versus low-power logic. In most designs, where the maximum performance of the FPGA is not required, the total power of a design can be reduced by about 50 percent or more.

Industry-Leading Process and Circuit Technology

The semiconductor industry is constantly battling the evolving challenges of small process dimensions through huge investments in equipment, process technologies, design tools, and circuit techniques. The challenge of increasing leakage power with small process geometries is felt industry-wide and a large number of widely used technologies at the 65-nm process node (and prior) are used to maintain or increase performance while managing leakage power. Altera continues to deliver leading-edge FPGAs using the latest industry capabilities as shown in Table 4.

Table 4. Altera Process and Design Techniques Adoption

Process or Design Technology	When Introduced by Altera	Benefit
All Copper Routing	150 nm	Increased performance
Low-K Dielectric	130 nm	Increased performance Reduced power
Multi-Threshold Transistors	90 nm	Reduced power
Variable Gate-Length Transistors	90 nm	Reduced power
Triple Gate Oxide (TGO)	65 nm	Reduced power
Super-Thin Gate Oxide	65 nm	Increased performance
Strained Silicon	65 nm	Increased performance

Copper Routing

Altera switched to an all-copper metallization for on-chip routing beginning with the 150-nm process node and used all-copper routing for all 130-nm, 90-nm, and 65-nm products, the earliest adoption in the FPGA industry. Copper replaced aluminum, providing reduced electrical and power resistance and thereby increasing performance.

Low-K Dielectric

A dielectric provides isolation between metal layers, enabling multiple routing layers. Moving to a low-k dielectric reduces the inter-routing layer capacitance, which significantly increases performance and reduces power. Altera was the first FPGA company to successfully adopt low-k process technology.

Multi-Threshold Transistors

The voltage threshold of a transistor affects the performance and leakage power of the transistor. Altera uses low-threshold voltages that produce high-speed transistors where performance is required and high-threshold voltages that produce slower, low-leakage transistors where performance is not required. Multi-threshold transistors are used in 90-nm and 65-nm Stratix series devices and 65-nm Cyclone series devices.

Variable Gate-Length Transistors

The gate length of a transistor affects its speed and sub-threshold leakage. As the length of a transistor approaches the minimum gate length of the 65-nm process, the sub-threshold leakage current increases significantly. Altera uses longer gate lengths to reduce leakage current in circuits where performance is not required. Where performance is critical, Altera uses short gate lengths to maximize performance. Altera has used variable gate lengths to reduce power in 90-nm and 65-nm Stratix series devices and 65-nm Cyclone series devices.

Triple Gate Oxide

The thickness of the gate oxide affects the performance and leakage current of a transistor. Altera uses three separate oxides (triple gate oxide) across the I/O circuitry and core logic. In Stratix III FPGAs, two of these core oxide thicknesses are used to enable low-performance transistors with minimum leakage, and high performance transistors for maximum performance.

Super-Thin Gate Oxide

The Stratix III triple gate oxide technology includes a super-thin gate oxide for high-performance transistors. These transistors enable the use of longer gate lengths while still maximizing performance, significantly reducing subthreshold leakage for a modest increase in gate-induced drain leakage and gate-direct tunneling leakage.

Strained Silicon

Strained silicon technology increases the transconductance of the transistor channel, thereby increasing the performance of the transistor. Altera uses strained silicon technology in Stratix III FPGAs for all transistors.

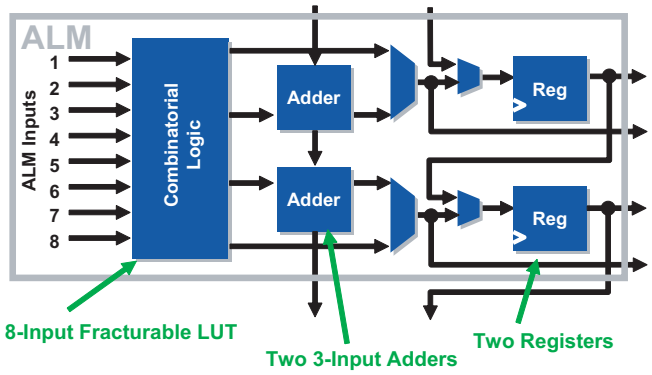
Designed for Lowest Power and Highest Performance

Altera has led high-end FPGA architecture innovation since the introduction of the first Stratix devices. Stratix III FPGAs leverage the first ALM logic architecture and MultiTrack interconnect fabric, which deliver the highest efficiency and performance compared to competing FPGAs.

Adaptive Logic Module

The adaptive logic module (ALM) technology introduced with Stratix II FPGAs maximizes performance and minimizes power by implementing 80 percent more logic functions than competitive architectures. Figure 10 shows the patented ALM architecture with the 8-input fracturable look-up table (LUT), two 2-bit adders, and two registers.

Figure 10. ALM Block Diagram



MultiTrack Interconnect

Stratix series devices also use the MultiTrack interconnect to maximize performance, minimize congestion, and minimize power. The MultiTrack interconnect provides the connectivity between different LABs and can be measured by the number of "hops" required to get from one LAB to another. Because adding interconnect hops increases capacitance, the fewer the hops, the less high-speed logic required to meet performance. As shown in Figure 11 and Table 5, the Stratix series MultiTrack interconnect provides the industry's best 1-hop interconnectivity, which yields the lowest possible power.

Figure 11. Stratix Series MultiTrack Interconnect

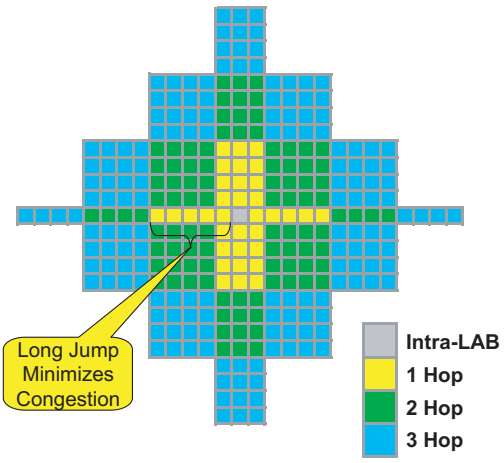


Table 5. Stratix Series Reachable LABs

Hops	Reachable LABs
1	34
2	96
3	160
Total	290

The combination of ALM and MultiTrack architectures enables more logic to be packed with less routing, thus increasing performance and reducing power.

Hierarchical Clocking

Stratix series FPGAs use hierarchical clocking to support up to 360 unique clocks. The propagation of every clock network can be controlled down to a LAB level. As part of the logic optimization in Quartus II software, logic with common clocks are grouped into LABs. Clocks are only propagated where the logic uses that clock. All other clock signals are shut down to minimize power consumption.

Figures 12 and 13 show before and after LAB clocking with placement optimization for low power. Figure 12 shows a pure performance-oriented placement that incurs increased clocking power. A more efficient grouping of clocks (Figure 13) minimizes clock power.

Figure 12. Hierarchical Clocking With Timing-Driven Placement

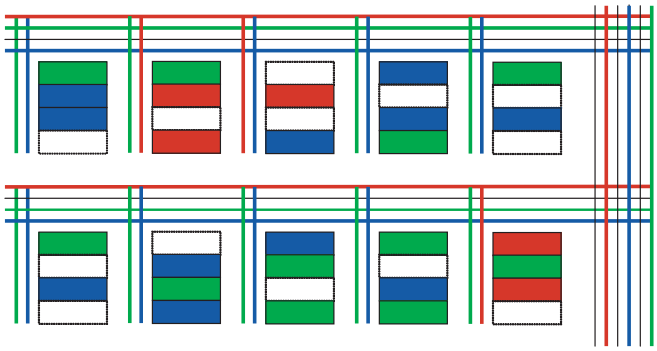
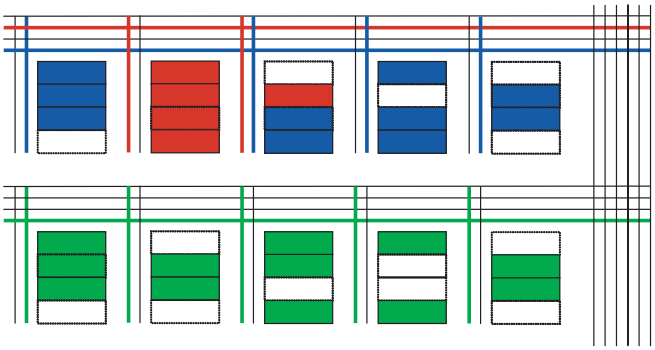


Figure 13. Hierarchical Clocking With Power-Driven Placement



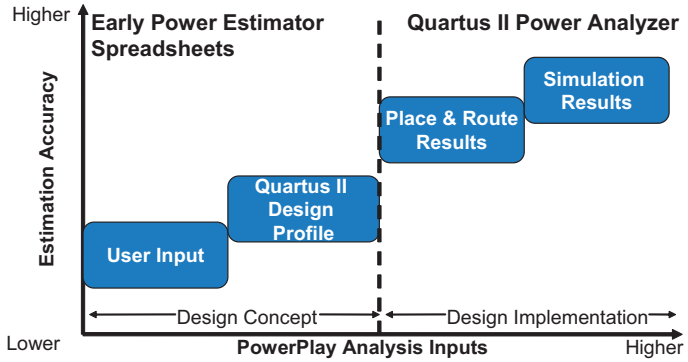
Software Programming Model

Taking advantage of the Stratix III low-power capabilities is seamless and automatic with Altera's Quartus II development software. This software has set the standard for FPGA power technology with fully automatic power optimization and the most accurate power estimation from any vendor.

Accuracy of Power Models

Altera supports power estimation from design concept through implementation, as shown in Figure 14. The designer uses the PowerPlay early power estimator (EPE) during the design concept phase and the PowerPlay power analyzer during the design implementation phase. These tools are the most accurate FPGA power analysis tools in the industry.

Figure 14. PowerPlay Analysis Tools-Accuracy vs. Implementation Detail



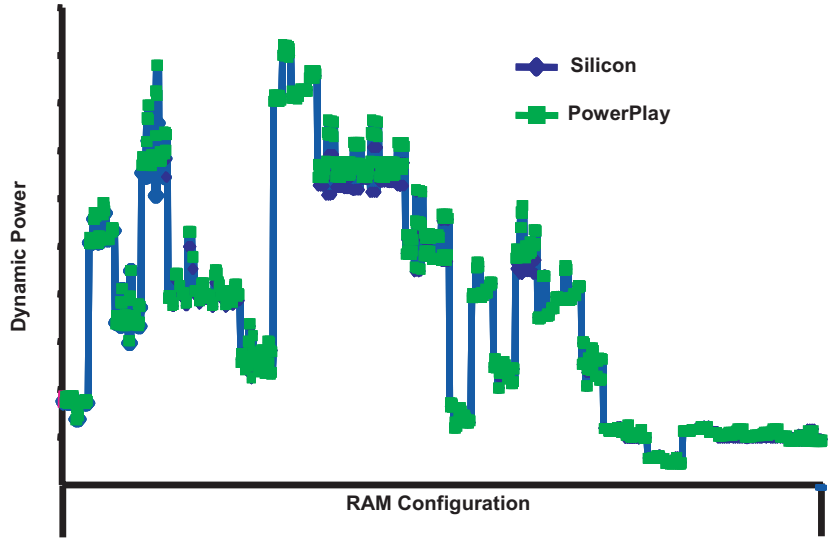
The PowerPlay EPE is a spreadsheet-based analysis tool that enables early power scoping based on device and package selection, operating conditions, and device utilization. The EPE has the industry's most accurate models of the functional components within the FPGA, but because the EPE is used before an RTL design is available, it lacks critical information such as logic configuration, placement, and routing, limiting its overall accuracy. Nevertheless, customers rely upon the EPE as their primary power estimation tool because it enables early design cycle estimates.

The PowerPlay power analyzer is a far more detailed power analysis tool that uses actual design placement and routing and logic configuration, and can use simulated waveforms to estimate dynamic power very accurately. The power analyzer, in aggregate, usually provides ± 10 percent accuracy when used with accurate design information.

Quartus II PowerPlay power models closely correlate to actual silicon measurements. Altera uses over 8,500 different test configurations to measure the power of individual components within a Stratix series device. Each configuration is focused on measuring a single circuit component of the FPGA in a specific configuration. Examples include DSP blocks in 9x9 mode, M9k memory blocks in x16 mode, and ALMs with specific logical configurations.

The test methodology is very straightforward and very accurate. The best way to accurately measure the power of a single block in a specific configuration in the FPGA is to configure the FPGA with all instances of a block measured in the configuration state under analysis. All other logic and functional blocks are configured for the lower power operating mode and are not stimulated. Then well-designed and repeatable stimulus patterns are run through all instances of the block being measured to generate an understood power profile. The resulting power consumed by the chip is largely the result of the large number of blocks under test, and the excess power can be subtracted from the total power. The resulting power, divided by the number of blocks configured, gives an accurate view of power for that mode of that block, as shown in Figure 15.

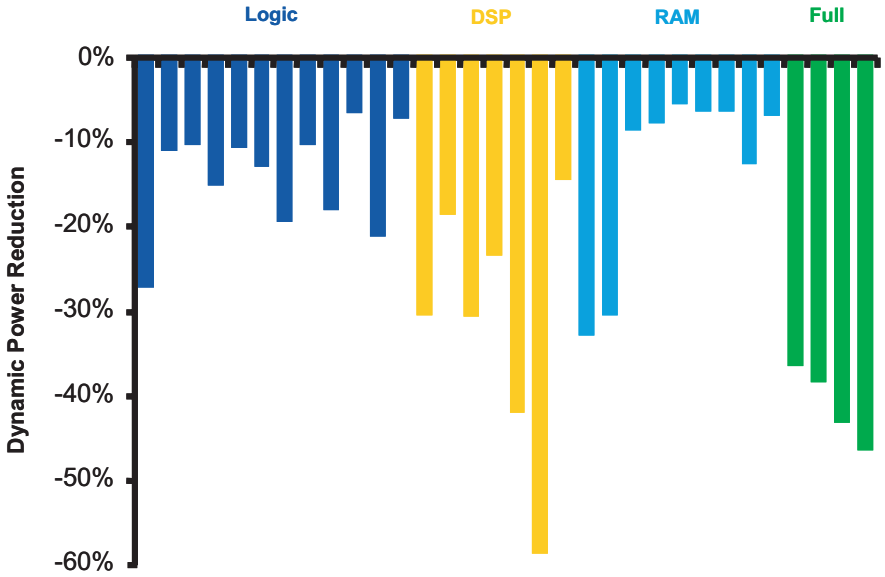
Figure 15. PowerPlay Power Estimate vs. Silicon Measurements for All RAM Configurations



Quartus II Power Optimization

Design implementation details can improve performance, minimize area, and reduce power. Historically, the performance and area trade-offs have been automated within the register transfer level (RTL) through the place-and-route design flow. Altera has taken a leadership position in bringing power optimization into the design flow, enabling power reductions of 10 to 40 percent over standard performance and area-optimized Stratix II designs (Figure 16). Quartus II PowerPlay optimization tools automatically use the new Stratix III architecture capabilities to reduce power further.

Figure 16. Stratix II Power Reductions Benchmarks Across Logic, DSP, RAM, and Balanced Resource Designs



Quartus II software has many automatic power optimizations that are transparent to the designer but provide optimal utilization of Stratix FPGA architecture details to minimize power, including:

- Optimizations in analysis and synthesis:
  - Transform major functional blocks
    - Map user RAMs so they use less power
  - Restructure logic to reduce dynamic power
    - Correctly select logic inputs to minimize capacitance on high-toggling nets
- Optimizations in fitter:
  - Reduce area and wiring demand for core logic to minimize dynamic power in routing
  - Modify placement to reduce clocking power
  - Trade speed for reduced power when routing non-timing-critical data signals
  - Set tiles with timing critical paths to high-speed mode and all other tiles to Low Power mode (Stratix III only)

Beyond 65 nm

Altera is actively developing next-generation product architectures based on 45-nm process technology through our partnership with TSMC, the industry-leading contract fab. Initial 45-nm test chips have already been fabricated. Altera is well positioned with industry-leading architectures, software, and next-generation process technologies to continue delivering the largest, highest performance FPGAs with lower power than competitive offerings.

Conclusion

While the move to very small process nodes delivers the expected Moore's law benefits of increased density and performance, the performance increases can result in significant increases in power consumption, introducing the risk of consuming unacceptable amounts of power. If no power-reduction strategies are employed, static power consumption can increase to critical levels. In addition, without any specific power optimization effort, dynamic power consumption can increase due to the increased logic capacity and higher switching frequencies that are attainable.

Altera consistently delivers leading-edge technology that maximizes performance and minimizes power. The Stratix III architecture, Programmable Power Technology, and selectable core voltage breakthroughs enable the lowest possible power for high-end FPGAs. In addition, Stratix III FPGAs continue Altera's practice of using the industry's best practices in process and circuit design to reduce power by 50 percent over previous-generation devices. The Quartus II design software offers the best power analysis and optimization in the entire FPGA industry. Overall, the Stratix III solution provides the performance designers need at the lowest possible power of any high-end FPGA.

Further Information

- "Standby and Active Leakage Current Control and Minimization in CMOS VLSI Circuits," Farzan Fallah and Massoud Pedram:
  http://atrak.usc.edu/~massoud/Papers/IEICE-leakage-review-journal.pdf
- More details on Stratix architecture advantages are available in *FPGA Architecture*:
  www.altera.com/literature/wp/wp-01003.pdf

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

Copyright © 2007 Altera Corporation. All rights reserved. Altera, The Programmable Solutions Company, the stylized Altera logo, specific device designations, and all other words and logos that are identified as trademarks and/or service marks are, unless noted otherwise, the trademarks and service marks of Altera Corporation in the U.S. and other countries. All other product or service names are the property of their respective holders. Altera products are protected under numerous U.S. and foreign patents and pending applications, maskwork rights, and copyrights. Altera warrants performance of its semiconductor products to current specifications in accordance with Altera's standard warranty, but reserves the right to make changes to any products and services at any time without notice. Altera assumes no responsibility or liability arising out of the application or use of any information, product, or service described herein except as expressly agreed to in writing by Altera Corporation. Altera customers are advised to obtain the latest version of device specifications before relying on any published information and before placing orders for products or services.