

Leveraging the 40-nm Process Node to Deliver the World's Most Advanced Custom Logic Devices

Introduction

Altera's launch of the Stratix® IV and HardCopy® IV device families in the second quarter of 2008 marked the introduction of the world's first 40-nm FPGAs and the industry's only risk-free path to 40-nm ASICs. For Altera, the event culminated over three years of exhaustive planning, development, and collaboration with its foundry partner—Taiwan Semiconductor Manufacturing Company (TSMC)—to deliver custom logic devices exhibiting uncompromised product leadership. Altera's subsequent announcement in the first quarter of 2009 of Arria® II GX and Stratix IV GT FPGA families results in the industry's most comprehensive transceiver-product portfolio. [Table 1](#) shows Altera's history of developing the world's first 40-nm FPGAs.

Table 1. Timeline for Development of Altera's 40-nm Devices

| Date | Milestone |
|---------|--|
| Q1 2005 | Altera begins developing 40-nm FPGA and HardCopy ASIC families, begins collaboration with TSMC on 40-nm process |
| Q4 2005 | Altera tapes out first of nine test chips for 40-nm devices |
| Q2 2006 | Evaluation of test chip structures |
| Q4 2007 | TSMC announces production-quality 45-nm process and stronger ties with Altera |
| Q1 2008 | TSMC announces 40-nm process |
| Q2 2008 | Altera announces the world's first 40-nm FPGAs, the Stratix IV device family, and first 40-nm HardCopy IV ASICs |
| Q1 2009 | Altera announces most comprehensive transceiver-product portfolio, including Arria II GX and Stratix IV GT FPGAs |

The 40-nm process node carries particular significance, as it provides a strong foundation for supporting Altera's leadership position in offering the highest performance, highest density, lowest power, and most cost-effective FPGAs and HardCopy ASICs.

Significance of the 40-nm Process Technology

The 40-nm process offers clear benefits over prior nodes, including the 65-nm node and the more recent 45-nm node. One of the most attractive benefits is higher integration, which enables semiconductor manufacturers to pack greater functionality into less physical space. The tangible results of this kind of density improvement have been reported at the International Electron Devices Meeting (IEDM) events, where leading semiconductor manufacturers present the results of their process technology efforts. The benchmark measurement is SRAM cell size, and [Table 2](#) shows SRAM cell sizes for recent process nodes reported at past IEDM meetings (listed in order of increasing cell size for 45-nm processes). As the table illustrates, process enhancements enable semiconductor manufacturers to deliver significantly greater functionality in less area.

Table 2. Smallest Reported SRAM Cell Sizes for 65- and 45-nm Process Nodes (1)

| Manufacturer/Alliance (2) | 65-nm SRAM (μm^2) | 45-nm SRAM (μm^2) | 32-nm SRAM (μm^2) |
|-----------------------------------|--------------------------------|--------------------------------|--------------------------------|
| TSMC | nr (3) | 0.242 | 0.15 |
| ST Micro, Freescale, NXP | nr | 0.25 | nr |
| Fujitsu | nr | 0.255 | nr |
| Intel | 0.57 | 0.346 | nr |
| IBM | 0.54 | 0.37 | nr |
| Texas Instruments | 0.49 | nr | nr |
| IBM, Chartered, Infineon, Samsung | 0.54 | nr | nr |
| IBM, Toshiba, Sony, AMD | 0.65 | nr | nr |

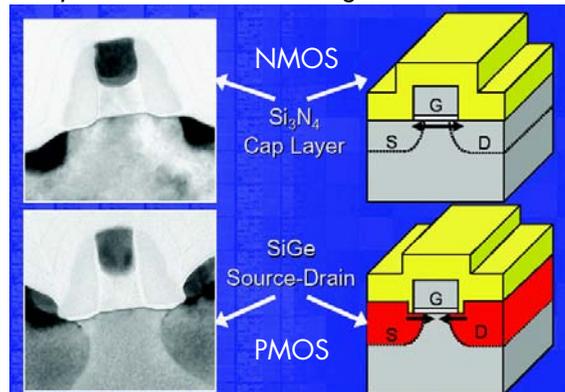
Note:

- (1) Source: Real World Technologies, "Process Technology Advancements at IEDM 2007"
- (2) Only companies/organizations reporting 65- or 45-nm SRAM cell size are shown.
- (3) nr = not reported

The 40-nm process also delivers clear performance benefits. The minimum transistor gate lengths of 40 nm are nearly 38.5 percent shorter than the gate lengths at 65 nm, and 11 percent shorter than the gate lengths at the 45-nm process. The corresponding lower resistance contributes to greater drive strengths at 40 nm, translating to higher performing transistors.

Altera achieves further performance gains with the use of strained silicon techniques. For example, Altera's devices benefit from tensile strain in NMOS transistors through a cap layer, and compressive strain for PMOS transistors through embedded silicon germanium in the source and drain (see Figure 1). These strained silicon techniques increase electron and hole mobility by up to 30 percent, and the resulting transistor performance is up to 40 percent higher.

Figure 1. Strained Silicon Techniques at 40 nm Enable Higher Performance Transistors



Although increased density and performance are valuable benefits, one of the most pressing design considerations for today's system developers is power consumption. The 40-nm node provides a benefit here, too, as smaller process geometries reduce the parasitic capacitances that drive up dynamic power consumption. Specifically, TSMC's 40-nm process technology provides active power downscaling of up to 15 percent over its 45-nm process technology.

Unfortunately, reductions in process geometry will also raise standby power unacceptably, if steps are not taken to reduce it. To address these and other growing power consumption concerns, Altera has taken aggressive steps to reduce both active and standby power in its 40-nm devices.

Combining the Leading Process and Device Architectures to Address Critical System Design Needs

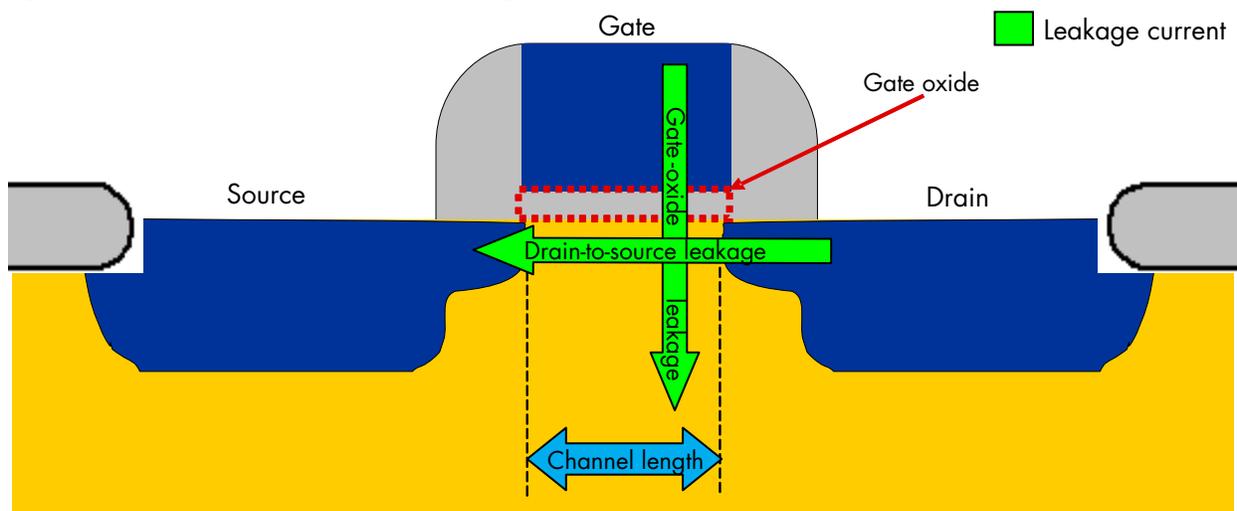
The move to the 40-nm node delivers the expected Moore's Law benefits of increased density and performance. Leveraging these process benefits and combining them with device architecture innovations enables Altera to continue offering the largest, highest performance custom logic devices in the industry. Accordingly, Altera® Stratix IV FPGAs and HardCopy IV ASICs deliver over 650K logic elements (LEs) and 13M ASIC gates, respectively. In the realm of performance, Altera's 40-nm device families can deliver over 600-MHz logic performance and transceiver performance of up to 8.5 Gbps, while maintaining the industry-leading LVDS I/O performance of up to 1.6 Gbps and single-ended I/O performance of up to 1066 Mbps, all without any compromise in signal integrity.

Besides the highest density and performance, Altera is also committed to delivering the lowest power consumption. The need for low power consumption is being driven today by the trends towards compactness of form factor, portability, and power efficiency. Product system enclosures are dramatically thinner and smaller, restricting airflow, heat sink size, and other thermal management solutions. Additionally, the energy component of the cost of operation is a top consideration for many applications, making low power consumption a significant competitive advantage, or in many cases a requirement. These shifts in design goals promote power consumption to the first-order selection criteria for system components.

FPGA vendors face growing challenges in managing power consumption as their devices grow in importance to occupy ever-increasing amounts of board functionality, in many cases expanding to implement the heart of the system. Balancing the demands for higher performance against the resulting higher power consumption is a significant effort.

At sub-micron geometries, semiconductor power consumption is a critical issue because static power can increase dramatically in the migration to more advanced processes. Smaller physical distances make it easier for current to leak. Both drain-to-source leakage and gate leakage are inversely proportional to channel length and gate oxide thickness, respectively, and can show dramatic increases as these lengths and thicknesses decrease (Figure 2).

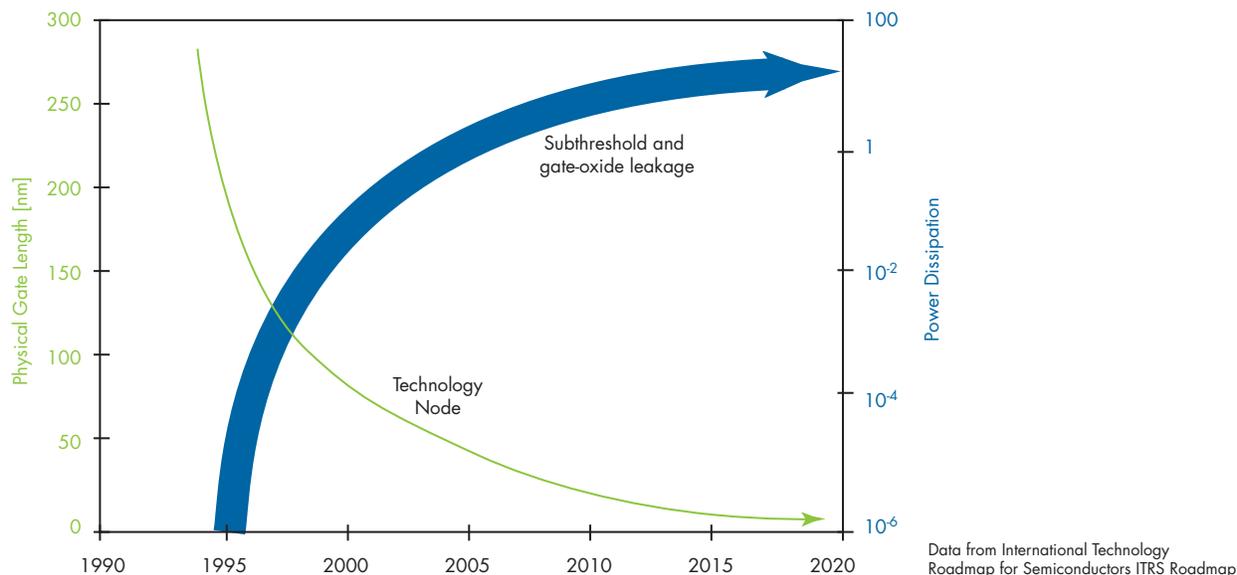
Figure 2. Transistor With Sources of Leakage Current



Source-to-drain leakage, also known as subthreshold leakage, is the dominant form of leakage. Here, current flows from the source to the transistor drain, even when the transistor gate is off. As transistors get smaller, it is harder to prevent this current from flowing, therefore the smaller 40-nm transistors tend to exhibit source-to-drain leakage with much greater magnitude than transistors on larger processes, all other parameters being equal.

The threshold voltage (V_t) of the transistor also influences the amount of source-to-drain leakage. The V_t of the transistor is the voltage at which the channel conducts current between the source and the drain. Small, high-speed transistors need a lower V_t to maintain the speed with which the transistor can be turned on and off via a gate control, but this increases the leakage because the transistor channel cannot be turned off completely. Another issue is gate oxide thickness, which—along with doping—influences V_t . A thinner gate oxide allows the transistor to be switched on and off faster, but it also allows greater leakage from the gate through the oxide to the substrate. These sources of leakage current increase as decreasing process geometries make smaller gate lengths possible, as shown in Figure 3.

Figure 3. Static Power Dissipation Increases Significantly at Smaller Process Geometries



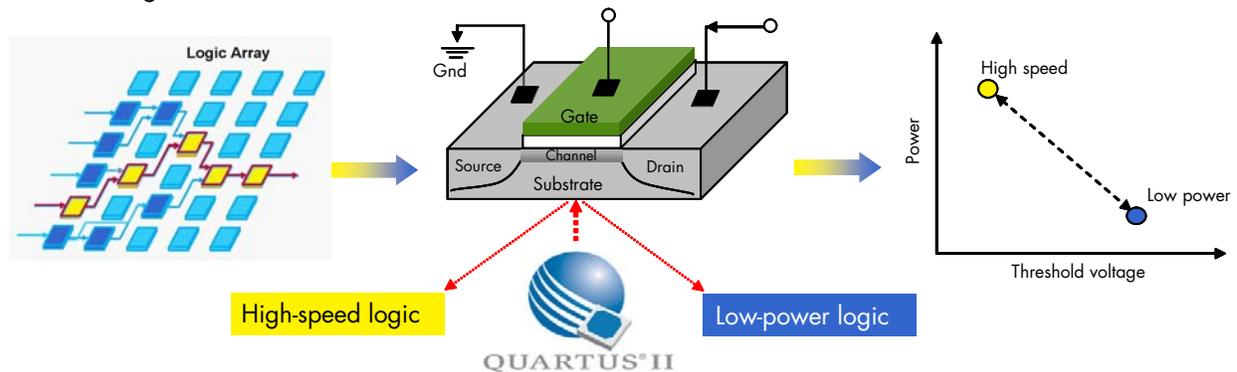
Altera primarily uses five methods to reduce leakage current, described in Table 3. All have a performance impact, slowing the transistor down from its maximum. However, Altera maintains overall device performance by making judicious choices to exchange performance for lower power at the transistor level. By analyzing circuit paths throughout the target device architecture, Altera IC designers identify where high-performance transistors can be best applied and where lower-performance transistors can be used to reduce leakage. In this way, the 40-nm process provides Altera IC designers with a platform to achieve the widest range of control and latitude for achieving both the highest performance targets and the most aggressive power consumption goals.

Table 3. Techniques Employed by Altera to Reduce Leakage Current

| Technique | Power Reduction Impact | Performance Impact |
|---|---|---|
| Increase transistor V_t via doping | Reduces source-to-drain leakage current | Raises voltage at which transistor turns on, reducing switching speed |
| Increase transistor channel length | Reduces source-to-drain leakage current | Reduces transistor switching speed |
| Apply thicker gate oxide | Reduces gate-to-substrate leakage current | Raises transistor V_t , reducing switching speed |
| Increase transistor V_t via Programmable Power Technology | Reduces source-to-drain leakage current | Raises voltage at which transistor turns on, reducing switching speed |
| Decrease V_{CC} | Reduces overall leakage current | Reduces switching speed |

In addition to the five methods described above, Altera applies its unique Programmable Power Technology to reduce static power. This patented feature, built into the silicon of Stratix IV devices, enables Quartus® II development software to change the transistor V_t in order to trade off performance and power based on the design requirements. Figure 4 shows a very high-level implementation of Programmable Power Technology, in which Quartus II software analyzes a user's FPGA design based on timing-driven compilation to select which transistors in the logic array should be in high-speed mode, and which should be in low-power mode. By changing the transistor V_t through the back bias voltage, the transistor is less leaky (increased V_t) in non-timing-critical paths, thus providing low power, but maintaining high performance where needed.

Figure 4. Programmable Power Technology (1) Allows Power/Performance Trade-Offs via Programmable Back-Biasing of Transistors



Note:

(1) This is a very simplistic “model” of Programmable Power Technology. Actual implementation varies and is patented.

To reduce the dynamic power consumption of its devices, Altera has lowered core voltage from the 1.1V used in prior device families to 0.9V in its 40-nm devices. The power consumed by a transistor during switching is proportional to V^2C (where C is capacitance), so reducing the supply voltage produces an almost quadratic reduction in dynamic power.

Lowering the core voltage also affects transistor performance, but Altera again leverages the higher performance of the 40-nm node to maintain high performance at the device level. As described earlier, Altera achieves much higher drive strengths in a given transistor at the 40-nm node compared to prior nodes, and its IC designers can trade off this drive strength for lower power consumption.

In addition, Altera enables the powering down of individual transmitter and receiver channels in its transceivers, which provide further power consumption reductions. Altera Stratix IV FPGAs also reduce active power consumption by offering dynamic on-chip termination (OCT). With dynamic OCT, the termination resistors in the Altera device can be turned on and off as needed. Turning off the termination resistors when they are not needed during memory read/write cycles eliminates the voltage drop across them, reducing power consumption by up to 1.2W with a 72-bit interface.

In total, Altera’s power reduction efforts with Arria II GX devices result in the lowest power FPGAs with 3.75-Gbps transceivers, which consume up to 65 percent less power than competing devices. In Stratix IV FPGAs, Altera’s power reduction efforts result in, on average, a 30 percent reduction in total (standby + dynamic) power consumption compared to similar designs implemented in its 65-nm Stratix III FPGAs.

From Technology Leadership to Smooth Production Ramps

Achieving the first 40-nm FPGAs is a significant event, but Altera’s goal extends beyond that to include maintaining the high quality and reliable delivery that it has demonstrated with products at prior process nodes. In this endeavor, Altera is well positioned to succeed due to its robust development practices, including a strong test chip plan, rigorous device checkout procedures, and a unique advantage to improving yields. All of these activities are reinforced and supported by the industry’s strongest foundry partnership.

Altera’s foundry partner, TSMC, has over 50 percent of the worldwide market share among dedicated foundries, as well as an annual research and development investment that is 55 percent greater than that of its nearest competitor. These investments have resulted in industry-leading positions in lithography and design-for-manufacturability (DFM) that further ensure TSMC’s success in delivering products at advanced process generations. Most importantly, at the 40-nm node, TSMC is the leader in immersion lithography, a process that combines lithographic lenses with clear liquids to preserve higher resolution light, enabling smaller, more densely packed devices. Immersion

lithography is the process of choice for the majority of semiconductor companies developing at the 45-nm node and below, and is acknowledged to be essential for the 32-nm node.

With TSMC, Altera actively maintains a dozen joint-process development teams addressing all aspects of process development, including power/performance, modeling, test chip planning, memory, reliability, poly fuse, DFM, RF/analog, ESD, and packaging. Each team has jointly agreed-upon deliverables and leaders from both companies, ensuring clear alignment and comprehensive follow-through.

Industry's Most Robust Test-Chip Practices

Altera has demonstrated with its 130-nm, 90-nm, and 65-nm devices that test chips are a valuable tool for early evaluation and refinement of architecture and device features on new semiconductor processes. This strategy helped Altera achieve smooth ramps to volume production of these devices, which has proved to be a point of distinction in the programmable logic industry. With the 40-nm node, Altera again is establishing a strong foundation for its latest generation of products with a robust plan of nine test chips.

This use of test chips represents a substantial investment due to the many mask sets involved. Altera's close collaboration with TSMC keeps the process efficient and minimizes cost. For example, TSMC runs numerous test wafers of its own to fully characterize and tune the fabrication methods, and then to monitor production. A close working relationship provides opportunities to "piggy-back" test structures on the foundry's wafers at the earliest stages, shortening the time-to-production for Altera's products and enabling its customers to gain access to the most advanced technology as soon as possible. In turn, Altera provides TSMC with opportunities to perform additional testing using its masks. Both companies benefit from the results.

By collecting and analyzing the test-chip data, Altera gains valuable insight into the impact of random and systematic variations, and is able to develop design strategies to reduce or eliminate them. Altera's significant investment in test chips ensures that customers are shielded from the many risks posed by leading-edge semiconductor design. This emphasis on risk management reflects Altera's commitment to deliver new technologies reliably without exposing customers to inconsistent or limited product availability or to products that fail to operate as specified.

Methodical Checkout Procedures

Beyond the test-chip stage, Altera performs a rigorous check, encompassing the development and manufacturing stages, to ensure that all of its silicon products operate exactly as specified. The checkout includes the following steps:

1. Altera's IC design team ensures that the design meets the functional, performance, and power specifications through a vast number of simulations, including statistical ones.
2. Through rigorous checking programs, Altera CAD and layout groups ensure that the implementation of the design fully meets all of Altera and TSMC's mask rules so the design can be processed successfully.
3. Cross-functional teams perform design-for-manufacturability (DFM) analysis on critical die areas to ensure robust manufacturing. This involves a detailed review of the design layout with the aim of removing any marginalities and optimizing the layout based on knowledge of the process technology to maximize yield.
4. TSMC ensures that the masks are manufactured properly. The resulting products can be manufactured in high volumes with no yield or functionality issues due to mask dimension marginalities or defects.
5. Altera works with TSMC to ensure that the silicon is manufactured properly, meeting all appropriate in-line physical specifications (layer thicknesses, line widths, etc.) and end-of-line electrical specifications (transistor characteristics, metal line resistances, etc.).
6. Altera's product engineers perform a full suite of characterizations at both the wafer level and the packaged unit level to ensure that the end product meets all specified functional, performance, and power specifications. They also characterize nonfunctioning units and work with other Altera teams to determine causes for yield loss, which is fed back to TSMC for yield improvement activity.

7. Altera's applications team tests the device from the user's point of view, exercising all device features, using Quartus II software to develop configuration files and program the device, testing I/O voltage levels, and verifying functionality of all architectural elements.
8. Altera's reliability group subjects both test chips and final products to rigorous environmental tests to ensure the short-term and long-term quality of the final product before it is shipped to customers.

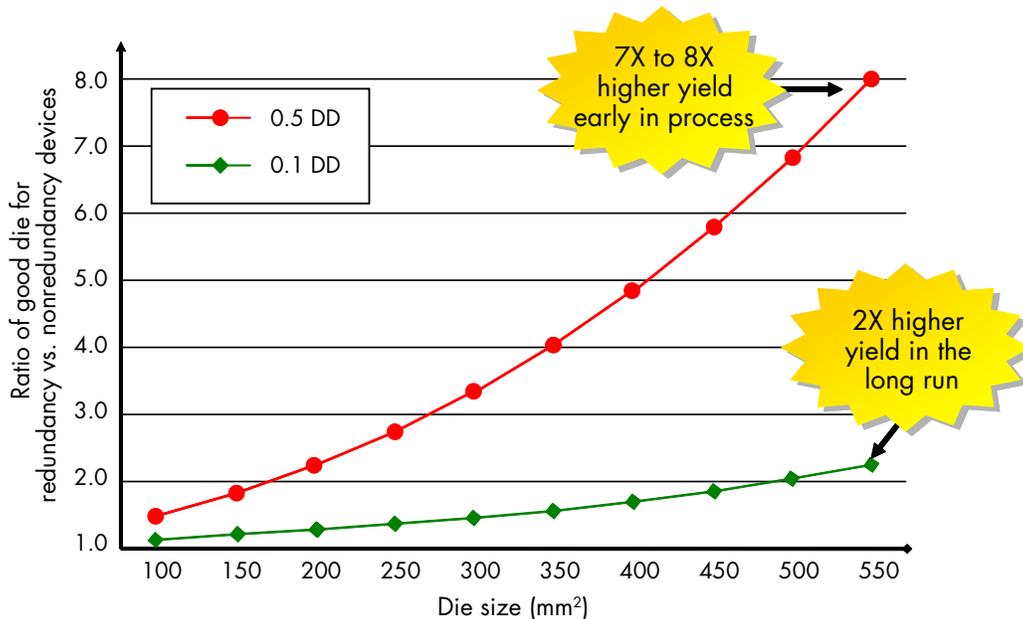
This uniform process is used and improved upon with each new device family offered by Altera. By applying these rigorous test and checkout procedures to every product, Altera ensures the highest levels of quality and reliability, as well as availability.

Unique Redundancy Technology Improves Device Yields

Altera is the only programmable logic vendor that leverages patented redundancy technology. Redundancy is a very effective method for improving device yields and device availability. Altera applies this technology by embedding extra, or "redundant," columns of circuitry into its FPGAs. If a column is subject to a manufacturing defect, it can be deactivated and the redundant column activated by the use of electrical fuses. This technology saves a die and thereby increases the total yield of a silicon wafer.

Redundancy is very effective with larger die, which are more likely to be affected by defects, especially in the early stages of a process or early in the life of a device. The addition of redundancy to the process improves yields for large die devices by up to eight times. In this way, redundancy improves product yields early in the process life cycle, brings the costs down more quickly, and increases overall availability. As the manufacturing process matures and defect densities improve, redundancy continues to play an important role, enabling Altera to achieve significant yield improvements in the long term (as shown in Figure 5). Overall, redundancy plays a major role in Altera's ability to achieve production-quality status for its products and reliable high-volume production more quickly than other programmable logic vendors, particularly with high-density products.

Figure 5. Redundancy Delivers Higher Yields Throughout a Product's Life Cycle



Supported by these practices, and indicated by its track record at previous nodes—all 90-nm devices delivered on schedule, and the world's first low-cost 65-nm FPGAs, Cyclone III family, delivered only three months after tape-out—Altera is well poised to deliver its 40-nm products reliably. Altera's track record at 65 nm also indicates a smooth ramp to production, as typified by Cyclone III FPGAs, which are manufactured in TSMC's two 300-mm GigaFabs less than a year after launch.

Conclusion

The 40-nm process comes with new design challenges to address, and the penalty for error is high. Mask costs grow about 50 percent each generation and for the 40-nm node run as high as \$3 million. Equally important, the cost of the design effort is growing because of increasing gate count and chip complexity, and growing more rapidly than mask cost. These barriers make 40-nm design prohibitive for all but a shrinking number of organizations.

However, Altera's business model enables it to make the heavy investments required to develop products using the most advanced semiconductor process and make them available off-the-shelf. Culminating a multi-year effort of planning, development, and close collaboration with the world's leading independent foundry, Altera's Arria II GX FPGA, Stratix IV FPGA, and HardCopy IV ASIC families enable early and broad access to 40-nm technology that would otherwise be out of reach. As a result, Altera customers gain access to the most advanced custom logic products delivering the capabilities, performance, density, and power consumption to address the most pressing needs of today's system designers.

Further Information

- *Altera at 40 nm: Jitter-, Signal Integrity-, Power-, and Process-Optimized Transceivers:*
www.altera.com/literature/wp/wp-01057-stratix-iv-jitter-signal-integrity-optimized-transceivers.pdf
- *40-nm Power Management and Advantages:*
www.altera.com/literature/wp/wp-01059-stratix-iv-40nm-power-management.pdf
- *Increasing Productivity With Quartus II Incremental Compilation:*
www.altera.com/literature/wp/wp-01062-quartus-ii-increasing-productivity-incremental-compilation.pdf

Acknowledgements

- Martin S. Won, Senior Member of Technical Staff, Customer Success Programs, Altera Corporation



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

Copyright © 2009 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.