

This paper outlines the performance and power requirements for next-generation programmable logic solutions to meet the demands of the growing Information and Communications Technology (ICT) sector. These requirements include:

- 2X higher performance compared to current solutions
- Up to 70 percent lower total power compared to current solutions
- I/O capabilities including high-speed serial transceivers supporting 56 Gbps as well as memory interface support options including DDR4 SDRAM and hybrid memory cube (HMC)
- Digital signal processing (DSP) performance over 10 trillion floating-point operations per second (teraFLOPS)

Introduction

The world has experienced dramatic increases in the demand for information bandwidth in recent years, and nearly every expectation is that this demand will continue to rise. In 2006, international Internet bandwidth was reported at 11 terabits per second (Tbps); by 2011 it has grown to almost 80 Tbps. Global bandwidth grew at a 57 percent CAGR during most of that period, with a 45 percent increase in 2011 alone. In 2013, almost 40 percent of humanity is estimated to be online, and by some estimates, the IP traffic from the world's consumers, industry, and governments will exceed a Zettabyte per year by the end of 2016, and at this level, the gigabyte equivalent of all movies ever made will cross global networks every three minutes ^{(1) (2) (3) (4)}.

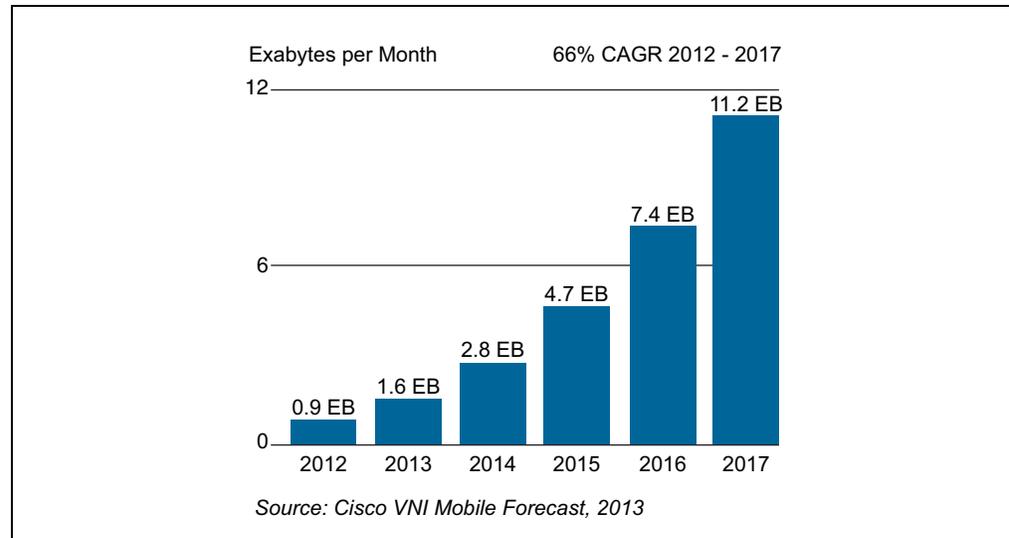
This seemingly insatiable demand for bandwidth is driving the ICT sector to overhaul the global communications infrastructure, from the data center servers that store and process the data, to the copper and optical networks that span the distances between physical locations, to the radio transmission towers that bridge the last mile to mobile devices. In turn, equipment manufacturers who supply the ICT sector are challenged to boost the capabilities of their products to process, package, frame, route, transmit, and receive more data via higher speeds, fatter pipes, and more multiplexing—or risk being left behind in the race to satisfy the global information appetite.

Delivering these capability boosts requires pushing the performance envelope, and examples of this need abound throughout the ICT infrastructure. Data centers place a high premium on the maximum processing power while minimizing cost, energy consumption, and physical footprint. In wired networking, while sales of 40 Gbps Ethernet (GbE) products are still growing, and the 100GbE market is nascent with the ratification of that standard in 2010, the IEEE has in 2013 realized the need to accelerate data rates an order of magnitude beyond the current mainstream, and

formally pursues a 400GbE standard. Mobile communications is also fueling this trend as a driver; in 2011, wired devices accounted for nearly 55 percent of IP traffic, but with the explosive growth of smart mobile devices, it is easy to envision that wireless devices will generate the majority. Figure 1 shows the projected growth in mobile data traffic, showing a 66 percent CAGR through 2017 ⁽⁴⁾.

With this growth in mind, mobile communications service providers are building out their Long-Term Evolution (LTE) networks to support this more data-intensive future where users expect the full processing power and connectivity of a wired device to accompany them wherever they go.

Figure 1. Growth in Mobile Data Traffic Through 2017



Higher Performance, Higher Power, and the Responsibility That Comes with It

If economics and environment were not a factor, the exponential growth in bandwidth might be accompanied by a similar increase in power consumption. However, all players involved in ICT expansion have incentives—economic and otherwise—from service providers, governments and end users, to improve efficiency and thus minimize the industry’s power consumption. As of 2012, the electricity consumption of the Internet is estimated at two percent of the world’s total, growing to potentially 10 percent by 2020. At a regional level, ICT is currently responsible for 5 - 10 percent of a typical economy’s total electrical consumption, and at a corporate level, ICT can account for up to 75 percent of all energy consumption. Data centers are acknowledged to be one of the fastest growing parts of the ICT sector—data center power demand grew 63 percent from 2011 to 2012, from 24 Gigawatts (GW) to 38 GW—and hence warrants special focus ^{(3) (5) (6) (7)}

From an environmental perspective, the carbon footprint of the ICT sector represents about 2 percent of worldwide greenhouse gas emissions, and is projected to grow at 6 percent CAGR. However, the influence of the ICT over emissions may extend far beyond that; the Smart 2020 Report finds that the ICT sector could help drive a reduction in emissions up to 15 percent in 2020 across the industries that rely on it through reporting, feedback mechanisms and other means. In economic terms, that reduction is estimated at almost \$950 billion in cost savings ^{(8) (9)}.

Reigning in the power of ICT equipment while boosting its capabilities is a daunting task, since power consumption of electronic systems rises when their designers leverage increased clock speeds, wider datapaths, or other common methods of increasing system performance. Worse yet, increased electronic product performance also results in more heat generated, and dissipating that heat and keeping the entire system within optimal operating temperatures consumes further power.

Fortunately, developers have options stemming from the advanced components available to architect their systems. For example, customization and flexibility at the hardware level are recognized as approaches to address these power challenges. In the case of data centers, Google customizes their servers to achieve high energy efficiency. This customization includes power supply selection, back-up batteries in the server racks, and matching server power with activity level. These efforts pay off in lower energy consumption; although Google is responsible for 2.8 percent of the world's volume of servers, it is responsible for less than one percent of the electricity used by the world's data centers ^{(10) (11)}.

Mobile communications offers another area where hardware customization can address the problem of growing energy consumption. The power associated with transmission signaling in mobile networks is nearly constant since it is provisioned to handle peak demand. However, actual traffic varies widely throughout the day and even drops off dramatically at “quieter” times such as night. Hardware solutions that are flexible and can be reconfigured to meet changing transmission needs on demand are recognized as a powerful approach to addressing this problem ⁽¹²⁾.

Equipment developers need the means to customize their hardware at the lowest levels to handle the opposing challenges of higher performance and minimizing power; this capability is delivered by components that are flexible enough to meet their exact requirements, which will optimize the power draw of their products.

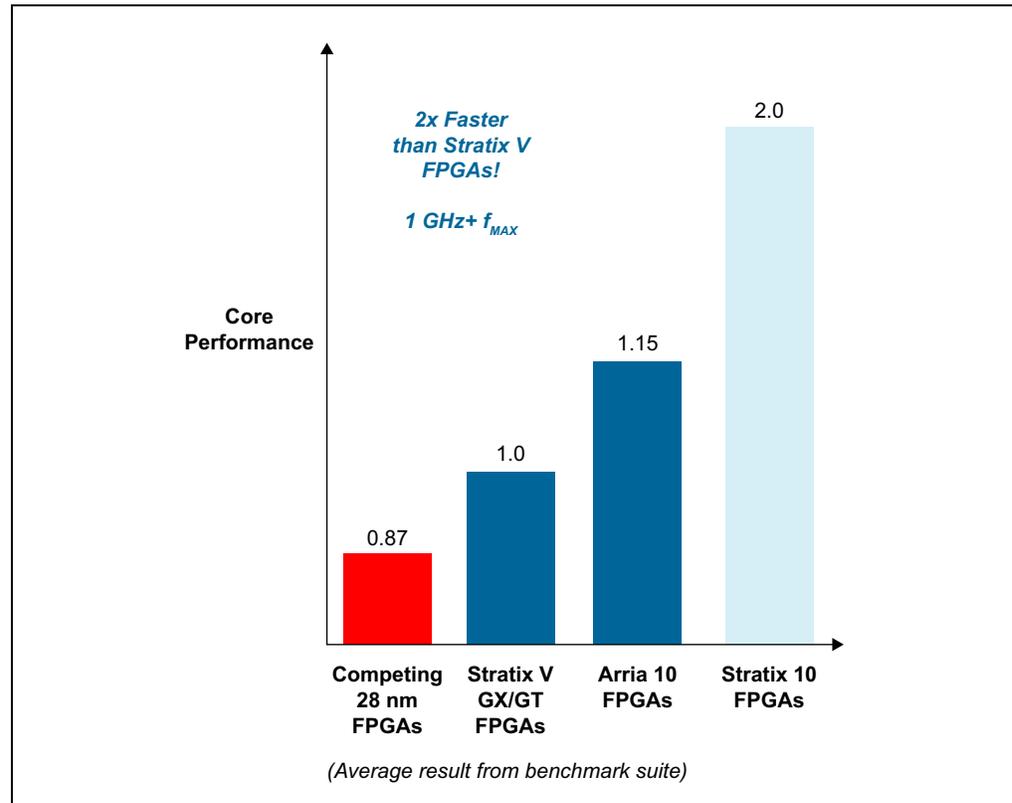
Generation 10 FPGAs and SoCs Address the Performance Challenge

To satisfy the needs of ICT equipment developers requiring higher performance, leading programmable logic vendors are exercising multiple product development strategies including leveraging the most advanced semiconductor processes, pushing high-speed serial (or transceiver) data rates to new highs, making architectural enhancements emphasizing performance, and supporting the latest memory interface standards HMC and DDR4 SDRAM.

For example, Altera is building its Stratix[®] 10 FPGAs and SoCs using the 14 nm Tri-Gate process from Intel, the most advanced Tri-Gate semiconductor process available ⁽¹³⁾. Leveraging this process advantage and a transceiver intellectual property (IP) portfolio and development team regarded by many as the strongest in the industry, Altera will deploy 56 gigabits per second (Gbps) transceivers with its Stratix 10 devices. At the architectural level, Stratix 10 devices also benefit from new floating-point DSP enhancements, enabling them to deliver DSP performance previously only possible with high-performance graphics processing units (GPUs).

With these improvements in place—along with a new ultra-performance architecture—Stratix 10 FPGAs will deliver Gigahertz fabric speeds, offering two times higher performance compared to the fastest FPGAs today, as shown in Figure 2, over 10 teraFLOPS of DSP performance, over 2.5 Tbps of serial memory interface bandwidth via HMC, and over 1.3 Tbps of parallel memory interface bandwidth via DDR4 SDRAM, making them the highest performance and highest bandwidth programmable devices.

Figure 2. Relative Performance of Stratix 10 Devices Compared to Other FPGAs



Tri-Gate Semiconductor Process Enables New Levels of Digital and Analog Performance

The Tri-Gate process represents a radical departure from prior semiconductor manufacturing technology by leveraging a new means of architecting a transistor that breaks through the power and performance limitations of traditional planar transistors. Tri-Gate transistors, also known as FinFET transistors in the industry, achieve these results with a 3-D structure that improves power and performance characteristics by addressing the shortcomings of current planar transistor architecture. Figure 3 shows a cross-sectional representation of a traditional transistor on the left, in which the gate is built in a single plane (thus the name “planar”). The effective width of the channel is shown in yellow. The channel width influences the transistor drive strength, with greater channel widths enabling higher drive strength and correspondingly higher performance. The Tri-Gate transistor, shown on the right, illustrates how the 3-D structure enables a greater channel width (also shown in yellow) without increasing the overall footprint of the transistor, resulting in higher performance without increased die area.

Figure 3. Cross Section of Planar and Tri-Gate Transistors Illustrating Structure Differences and Resulting Performance Improvements

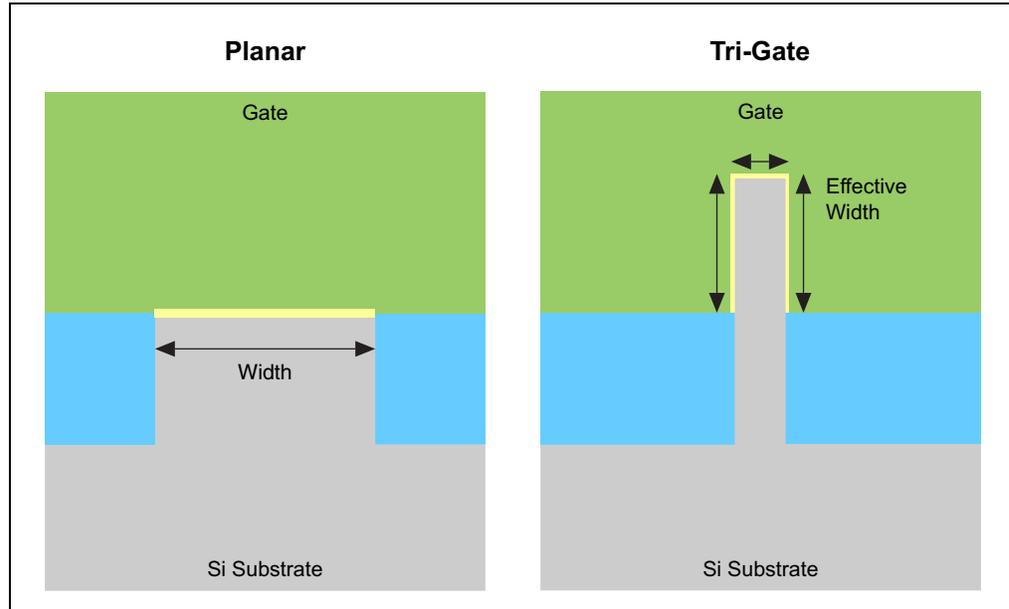


Figure 3 also shows how Tri-Gate transistor performance can be further improved with addition of more vertical structures (or “fins”) controlled by the same gate, or by increasing fin height, both of which increase the channel width without increasing the footprint of the transistor.

In addition to improving the performance of digital logic, Tri-Gate transistors provide performance benefits for analog circuits such as high-speed transceivers. These improvements stem from the increased gain that Tri-Gate transistors deliver relative to planar transistors on the most advanced submicron semiconductor processes. Gain is an essential element in analog circuit design, providing the following key benefits:

- Better noise rejection due to higher output impedance (R_{OUT})
- Reduced deterministic offset in amplifiers
- Better precision in amplifiers via the ability to amplify smaller signals
- More drive strength due to increased transconductance (g_m)
- Increased f_{MAX} due to increased transconductance

Figure 4 shows how transistor gain has been trending across recent submicron processes, where the gain is expressed as the product of transconductance and transistor output impedance. Transconductance in FET transistors represents the amount of change in drain current as a function of the change in voltage between gate and source, and a high value is desirable, representing the ability to drive large currents with small changes in gate voltage. Transistor output impedance is the impedance seen by the circuit that is driven by the transistor, and a high value results in better isolation for operational amplifiers (op-amps) and regulators, which reduces sensitivity to noise and therefore reduces jitter.

Figure 4. Tri-Gate Transistors Provide Improved Analog Performance (Intel Developer Forum, June 2012)

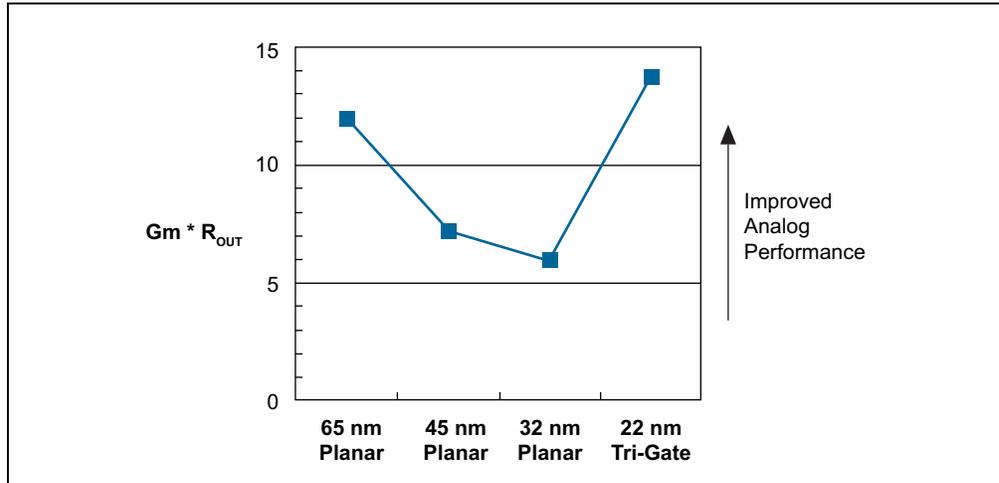


Figure 4 also shows that transistor gain has been trending lower in planar transistors on recent generations of semiconductor processes. This reduction in gain makes it harder to develop high-speed transceivers. To achieve the same performance, longer gate lengths are required for higher gain, which increases the loading and thereby increases both power and die area. Lower gain also results in higher offset error, which in turn requires more calibration and adaption circuits to correct, leading to increases in power and die area. Figure 4 shows a dramatic increase in gain resulting from 22 nm Tri-Gate transistors compared to prior generation planar transistors.

The increase in gain that derives from Tri-Gate transistors is a key enabler to delivering the 56 Gbps transceiver data rates that will be supported by Stratix 10 devices. Since the 14 nm Tri-Gate process is the second-generation process for Intel, it also represents a process shrink relative to the 22 nm process and any first-generation FinFET processes, which further reduces area and power consumption. The result is that not only will 56 Gbps data rates be made possible for the first time in any programmable logic device; they will also be the lowest power consumption possible for their level of signal integrity quality.

Industry's First in Performance: Hard IP for Floating-Point DSP and the Highest Levels of Support for HMC and DDR4 SDRAM

Altera is also making improvements to its award-winning variable-precision DSP blocks, adding structures that are optimized for single-precision floating-point operations. With this enhancement, Stratix 10 devices will offer over 10 teraFLOPS, a performance level that has never been offered in any off-the-shelf device. Coupled with this DSP improvement is corresponding increases in memory bandwidth stemming from HMC support and DDR4 memory support. For HMC support, a single Stratix 10 device can support up to 32 x 4 lanes at 15 Gbps each for a maximum bandwidth (including transceiver (TX) and receiver (RX)) of over 2.5 Tbps, assuming an efficiency of 67 percent. Likewise, a single Stratix 10 device can support up to six

72 bit wide DDR4 SDRAM interfaces at 3.2 Gbps each, for a total of 1.382 Tbps of memory bandwidth. [Table 1](#) summarizes the key performance improvements of Stratix 10 FPGAs over the prior generation of Stratix V FPGAs, demonstrating a performance gain on the order of two times or better across major device characteristics.

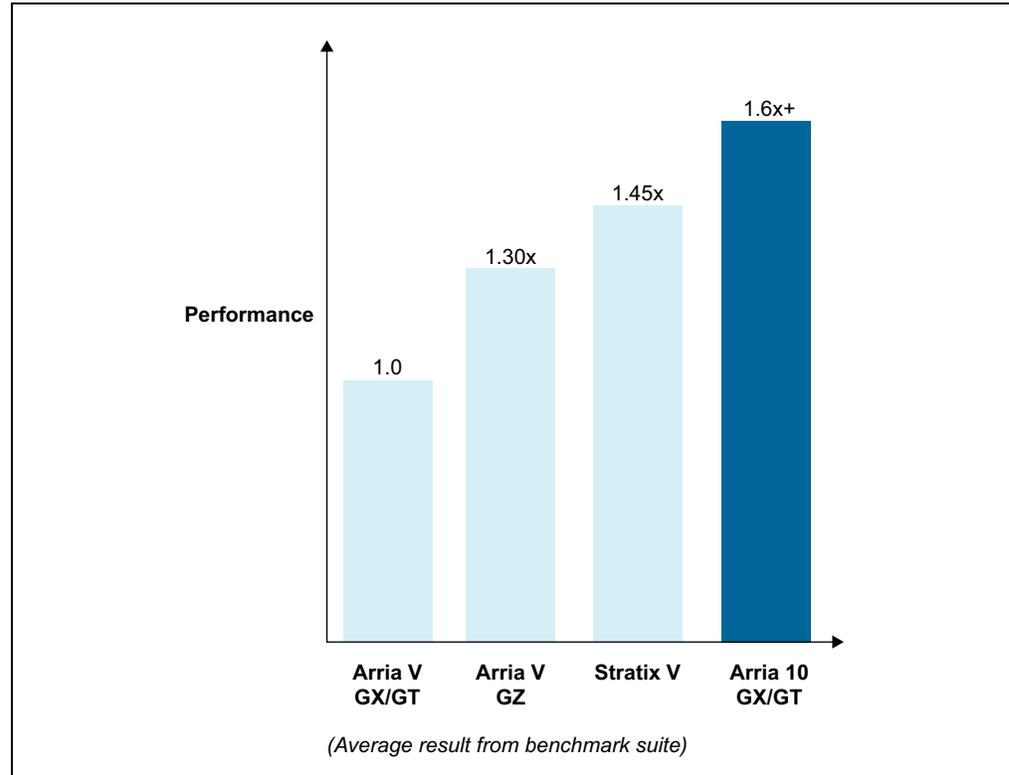
Table 1. Summary of Stratix 10 Device Performance Improvements vs. Prior-Generation Stratix V FPGAs

	Stratix V FPGAs	Stratix 10 FPGAs	Advantage
Density	1,000 K LEs	4,000 K LEs	4x
TeraFLOPS	1	10+	10x+
Core Performance	500 MHz	1 GHz+	2x
Transceiver Maximum Data Rate	28 Gbps	56 Gbps	2x
DDR Memory Interface Rate	1,866 Mbps	3,200 Mbps	1.7x
<i>2x Performance Advantage with Stratix 10 FPGAs and SoCs</i>			

Performance Improvements in Arria 10 FPGAs and SoCs

Arria® 10 FPGAs and SoCs comprise the midrange portion of Altera’s Generation 10 product portfolio, and they feature performance levels that have never been available in any midrange FPGA. Arria 10 devices are 60 percent faster than the prior-generation Arria V FPGAs, and 15 percent faster than Stratix V FPGAs, as shown in [Figure 5](#). Stratix V FPGAs are the fastest FPGAs of the 28 nm product generation, delivering a full speed grade higher performance compared to the fastest 28 nm competitor. By delivering higher performance than the fastest existing FPGAs while consuming lower power, Arria 10 devices provide an ideal power-reduction path for the current generation of high-end and midrange 28 nm FPGA designs.

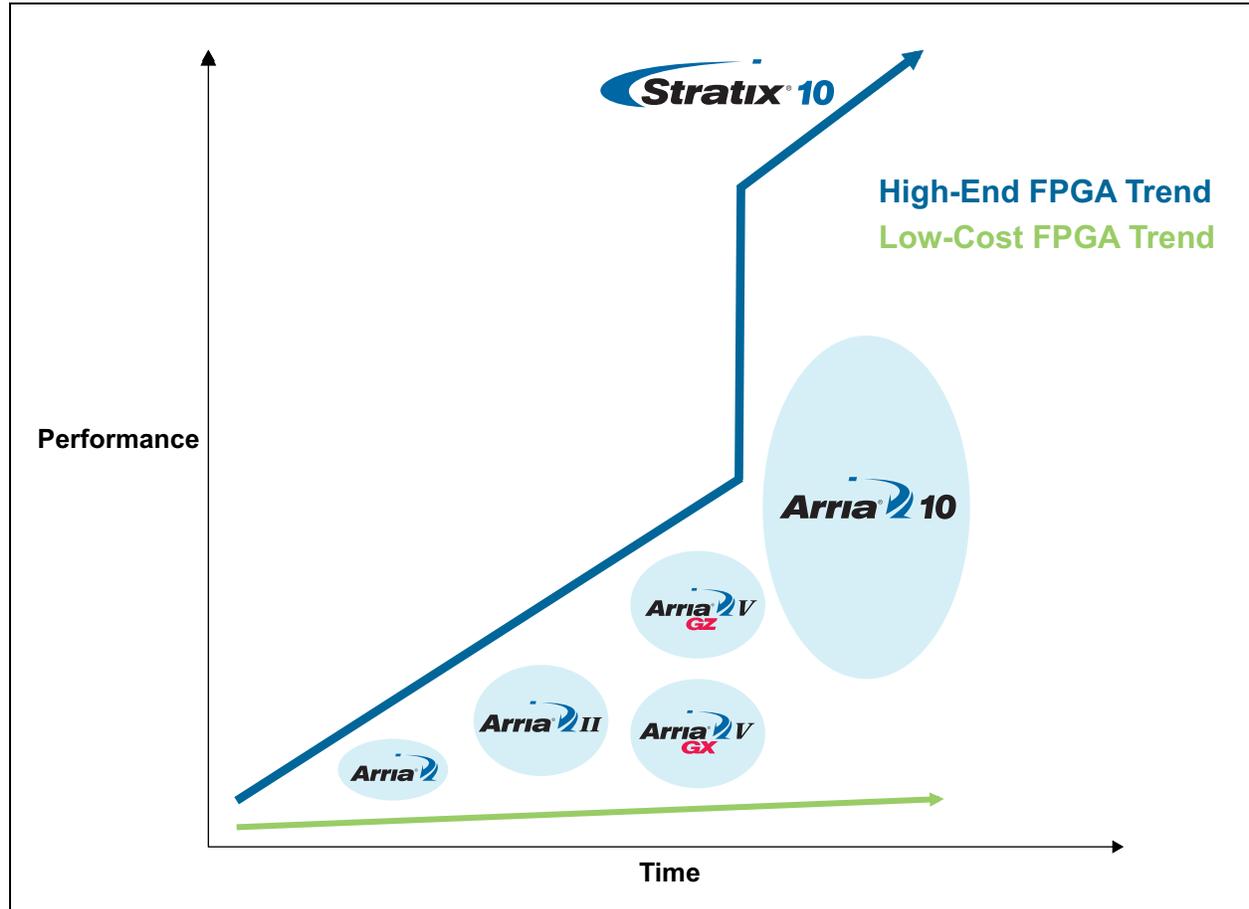
Figure 5. Arria 10 FPGA and SoC Performance Levels Compared to Prior-Generation FPGA Families



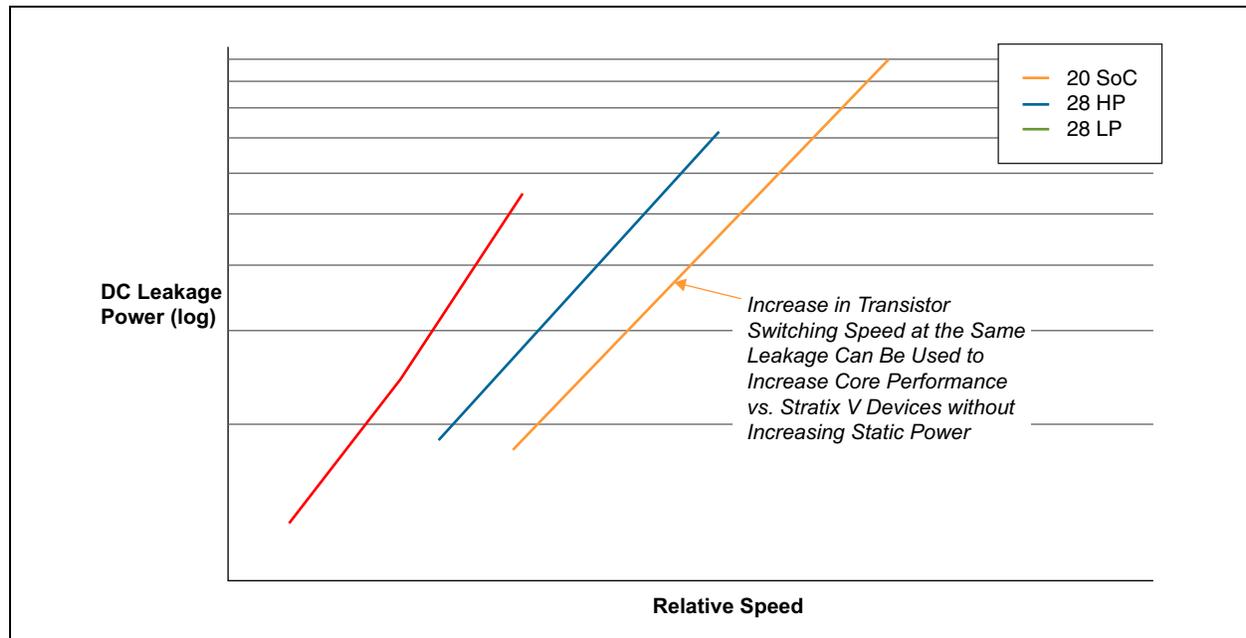
This marked increase in performance for Arria 10 devices is a dramatic shift from prior generation FPGA midrange products. Altera created the midrange FPGA category in 2007 with the first Arria GX FPGAs, when the performance gap between high-end FPGAs such as the Stratix device family and low-cost FPGAs like the Cyclone® device family was wide enough such that a clear market need for another product category between them went unaddressed, as shown in [Figure 6](#).

[Figure 6](#) illustrates another reality about the midrange product space. As time goes on, the gap between high-end and low-cost widens, forcing programmable logic vendors to choose how they will address the space with their various product offerings. The blue areas marked Arria GX, Arria II, and Arria V FPGAs show how Altera has addressed this gap over the past three generations. [Figure 6](#) also shows how the coming product generation includes the singular event of the introduction of 14 nm Tri-Gate FPGAs, and the resulting step increase in high-end FPGA and SoC performance. To address this abrupt widening of the gap, Altera has elevated the capabilities and performance of its Arria 10 FPGAs and SoCs far beyond what is expected of a midrange FPGA, and also broadened its density range, reinventing the definition of a midrange FPGA and SoC.

Figure 6. Midrange FPGA Category Created to Address Widening Performance Gap Between High-End and Low-Cost FPGAs



The increased performance levels of Arria 10 devices arise from three sources— the 20SoC semiconductor process, architectural improvements, and support for higher transceiver and I/O interface rates previously found only in high-end FPGAs. Figure 7 illustrates the performance benefit stemming from the 20SoC process. Figure 7 also shows that at the same leakage power level, the 20SoC process offers transistors that have higher switching speed than the 28HP process. This means the semiconductor devices built with the 20SoC process can deliver higher performance at the same leakage power, and by judicious selection of transistors, the resulting 20SoC-based devices can offer higher performance at lower leakage power, as is the case with Arria 10 devices.

Figure 7. Transistor Switching Speed in 28HP, 28LP, and 20SoC Processes Relationship to Leakage Power

Arria 10 devices also benefit from architecture enhancements to improve their performance. Many routing paths in Arria 10 devices have been accelerated relative to their counterparts in the prior architecture. The hard memory controllers have also been enhanced to support higher memory interface rates and also wider interfaces up to 144 bits. Transceiver data rates have likewise been improved, and the Arria 10 devices will feature the industry's first 20 nm transceivers operating at 28 Gbps data rates. Altera demonstrated the world's first 20 nm transceivers operating at 32 Gbps in April of 2013, providing early validation of its industry-leading transceiver capability. These transceivers will also provide the world's first HMC support in a midrange FPGA, enabling Arria 10 FPGAs to support up to 1.2 Tbps of serial memory bandwidth in a single device. [Table 2](#) summarizes the performance improvements in Arria 10 compared to its prior generation Arria FPGAs, demonstrating a performance gain on the order of two times or better across major device characteristics.

Table 2. Summary of the Arria 10 Device Performance vs. Prior-Generation Arria V GX/GT Devices

	Arria V GX/GT FPGAs	Arria 10 FPGAs	Advantage
Density	500 K LEs	1,150 K LEs	2x
Multipliers	2,312	3,356	1.5x
Core Performance	300+ MHz	500+ MHz	1.6x
Transceiver Maximum Data Rate	10 Gbps	28 Gbps	2.8x
1,333 Mbps	2,666 Mbps	2x	1.7x
2x Performance Gain with Arria 10 FPGAs and SoCs			

With Great Performance Comes Great Power

The next generation of FPGAs and SoCs stands poised to benefit from the very latest semiconductor technology to deliver high performance. However, the successive process shrinks that drive the increased integration and capability in semiconductors have also presented the critical problem of increasing leakage power. The increase in leakage power that has accompanied each process advance has been steadily undermining the value of each process shrink. The leakage power issue is so serious that in its 2009 report, the International Technology Roadmap for Semiconductors (ITRS) describes the situation in terms of an existential crisis:

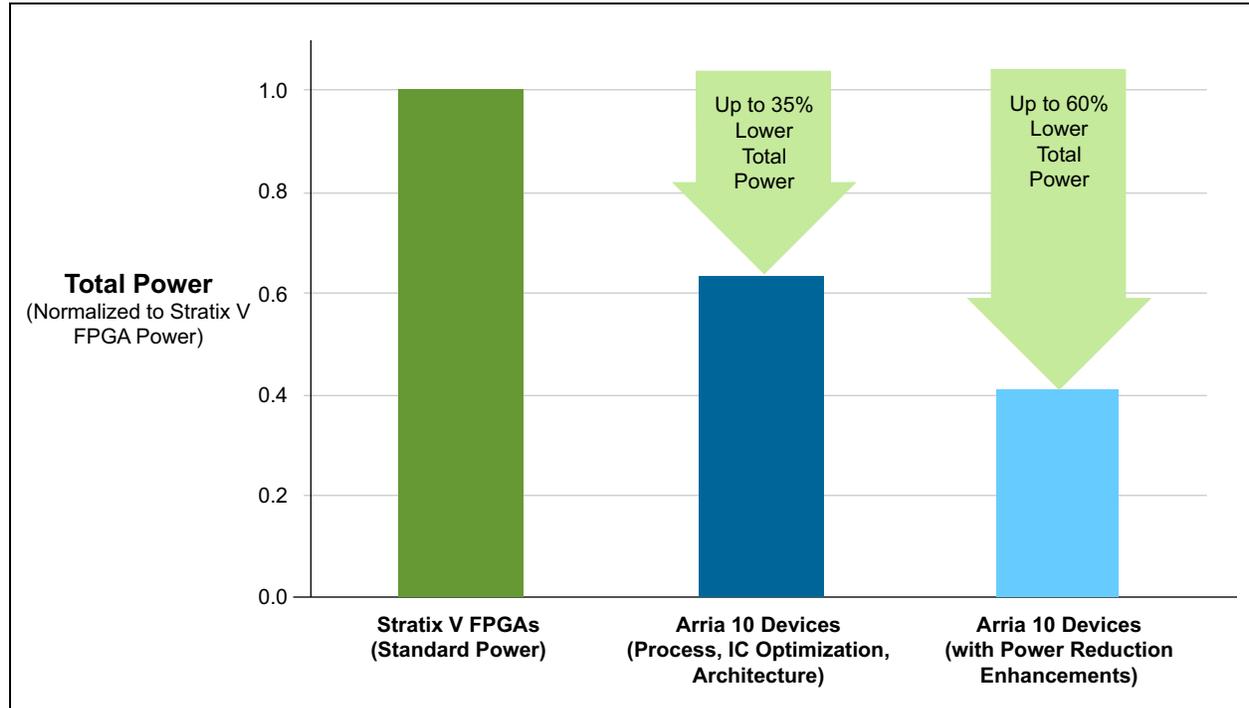
While power consumption is an urgent challenge, its leakage or static component will become a major industry crisis in the long term, threatening the survival of CMOS technology itself, just as bipolar technology was threatened and eventually disposed of decades ago (14).

The magnitude of this problem has given rise to significant effort to address it, resulting in major advances challenging the traditional CMOS transistor architecture that has stood for half a century. These innovations will fuel new capabilities in semiconductors and in particular programmable logic, as Altera will be among the first to leverage them in its next-generation product portfolio.

Power Reduction in Next-Generation FPGAs and SoCs

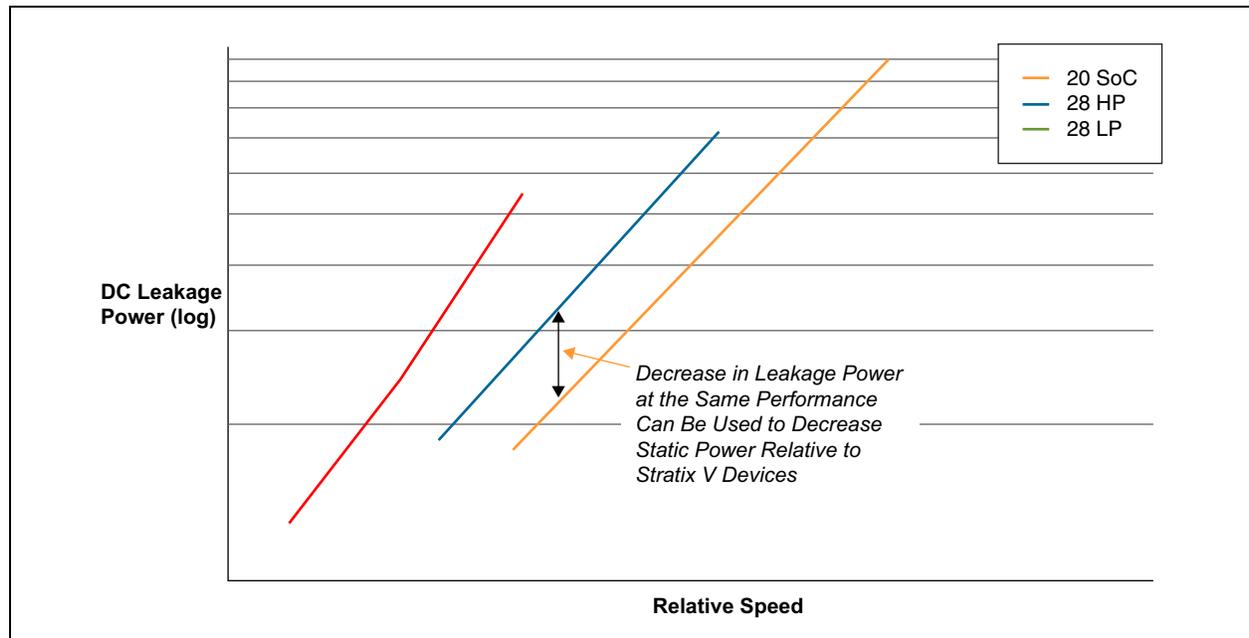
Arria 10 devices demonstrate significant power reductions compared to prior generation FPGAs, owing to their 20 nm process, architecture, and a comprehensive set of power reduction features. When compared to prior-generation 28 nm FPGAs, Arria 10 devices consume up to 60 percent lower total power. [Figure 8](#) illustrates this power reduction. The left bar shows the average power consumption of a set of target applications implemented in 28 nm Stratix V FPGAs. The dark blue bar in the middle shows the power reduction when using Arria 10 devices due to process, architecture, and voltage differences. For example, the 20 nm process offers a geometry shrink compared to the 28 nm process, which reduces capacitance and thereby lowers dynamic power. The final bar on the right side of [Figure 8](#) shows the potential total power of designs implemented in Arria 10 devices when all of its power reduction features are applied to their maximum effect. These power reduction features can reduce both core dynamic and core static power, resulting in designs that consume up to 60 percent less power when compared to the same design implemented in a 28 nm Stratix V FPGA.

Figure 8. Total Power Reduction in Altera’s 20 nm Devices (Arria 10 FPGAs and SoCs) Compared to Prior-Generation 28 nm FPGAs (Stratix V Devices)



Arria 10 also features more hard silicon IP, such as memory controllers, PHYs, and forward error correction (FEC) that also reduces capacitance and power compared to the same functions implemented in the programmable logic fabric. In addition, Arria 10 devices use lower voltages in comparison to 28 nm FPGAs. Finally, the process that Altera is employing to build Arria 10—the 20SoC process from TSMC—offers some transistors that consume less leakage current than the 28HP process that is used for Stratix V FPGAs, which reduces static power.

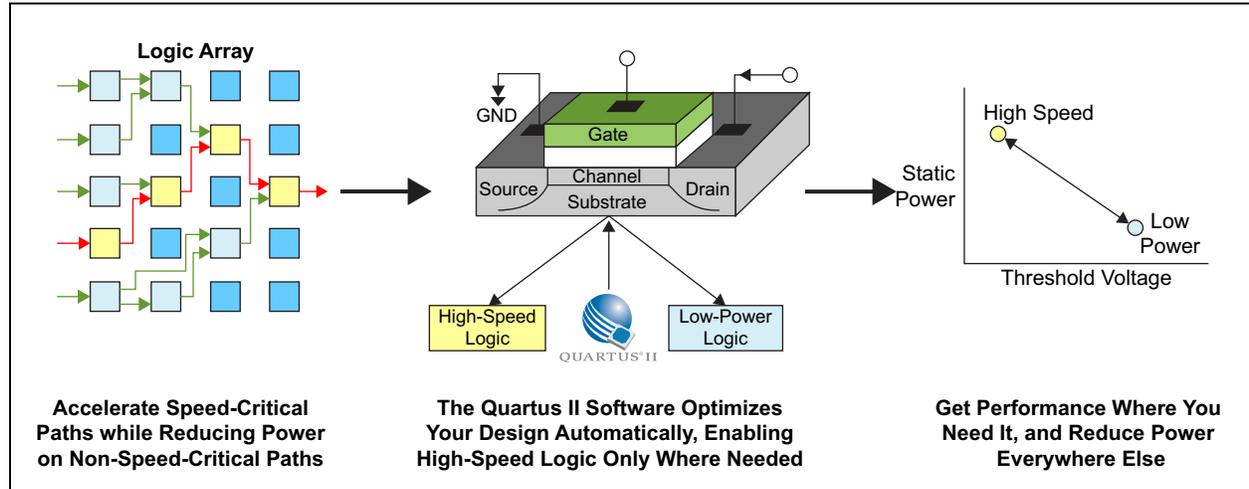
The relationship between the 28HP process, the 20SoC process and their leakage power characteristics is shown in [Figure 9](#). Leakage power is shown on the vertical axis, and transistor switching speed is shown on the horizontal axis. The blue line indicates transistors available within the 28HP process, and the orange line similarly shows transistors available for use in the 20SoC process. IC designers can choose from the transistor devices available within the chosen process to build their devices. The lower placement of the orange line relative to the blue line shows that for the same switching speed, the transistors in the 20SoC process consumes less leakage power.

Figure 9. Transistor Switching Speed in 28HP, 28LP, and 20SoC Processes Relationship to Leakage Power

Comprehensive Suite of Power Reduction Features

Arria 10 devices offer the most power reduction features of any high-density FPGA. The offering begins with programmable power technology (PPT), an Altera-patented method of tuning the switching speed of logic elements in the speed-critical paths of a user's design illustrated in [Figure 10](#). With this method, Altera's Quartus® II software automatically identifies the speed-critical paths in a user design and places the associated logic into a high-speed mode. This tuning is accomplished through changing the transistor threshold voltages in the path to a lower value, which increases its switching speed. All other transistors in the logic fabric that do not require the highest speed are tuned to a higher threshold voltage, which reduces their static power consumption. With PPT, the static power of the logic fabric or device core can be reduced by up to 20 percent.

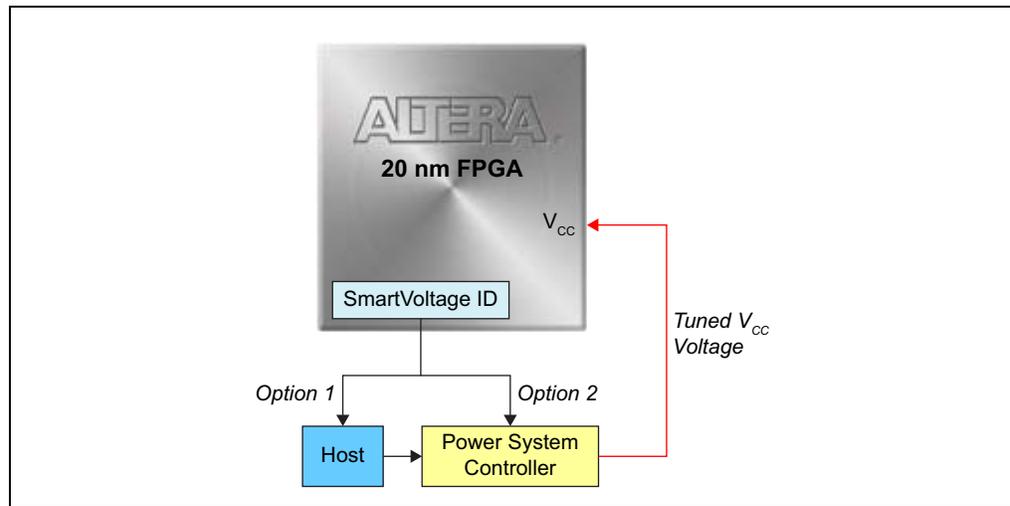
Figure 10. Altera's Patented PPT Reduces Static Power by Placing Speed-Critical Paths into High-Performance Mode and Leaving Other Logic in Power-Conserving Mode



SmartVoltage ID

The second Arria 10 power reduction feature is SmartVoltage ID. SmartVoltage ID leverages the variability of the semiconductor process to enable lower voltage operation while still retaining their specified performance level, a feature similar to one now found in many microprocessors and graphics processors. Semiconductor process variability means that some devices across a given distribution are faster, while others are slower. Altera guarantees that its devices will meet specified minimum levels of performance when powered at the standard V_{CC} voltage level. The distribution means that some of the devices are faster than the specification when powered at the standard V_{CC} voltage level. These faster devices can be powered at a lower-than-standard V_{CC} voltage level, and still meet the minimum specified performance.

During the manufacturing process, Altera tests all Arria 10 devices and programs each one with a SmartVoltage ID, which indicates the lowest V_{CC} voltage at which it can run while still meeting performance specifications. Arria 10 device users can then choose to use the SmartVoltage ID to reduce the power of their designs. Figure 11 shows how the SmartVoltage ID is used in a system to reduce power. The Arria 10 device powers up at the standard V_{CC} voltage level. The SmartVoltage ID level is then read from the Arria 10 device, either by an intelligent host, or by a power system controller that has this capability. This capability is featured in many power regulator products, driven by processors and GPUs.

Figure 11. Applying SmartVoltage ID in System

Arria 10 devices support industry-standard methods to communicate the SmartVoltage ID, including I2C and PMBus. The power regulator then applies the tuned lower voltage corresponding to the SmartVoltage ID, resulting in lower power consumption.

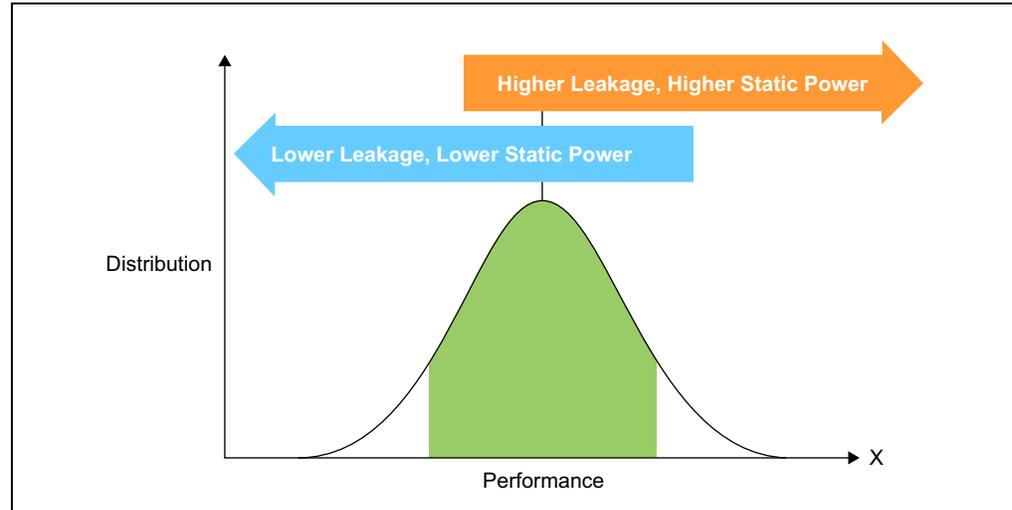
With the current implementation of SmartVoltage ID, Altera guarantees a maximum static power that is up to 40 percent lower compared to the worst-case static power of devices where SmartVoltage ID feature is not used. This lower static power results in power savings to users since they must design their power regulators to handle the worst case. There is also a dynamic power reduction associated with using SmartVoltage ID, which is apparent across a distribution of devices, but is not guaranteed for any single device because the actual SmartVoltage ID value can vary from one device to another.

Consider a million-logic element Arria 10 device running a high-performance design, in which the V_{CC} rail consumes 40 amps. At the standard V_{CC} voltage of 0.9 V, this result in 36 W of power consumed. When SmartVoltage ID is used, the actual V_{CC} voltage of this device could be 0.86 V, which results in 34.4 W. Across a year of operation, the savings is over 14 kilowatt hours (kWh) for the single device. Multiplied across potentially thousands of devices or more per system, or per user installation, the Opex savings can be significant.

Lower Static Power Device Versions

Another power reduction option for Arria 10 FPGAs and SoCs is the availability of versions that consume lower static power. Altera can offer these devices because of the distribution due to the semiconductor manufacturing process. Across any given manufacturing distribution, there will be devices that draw higher leakage current and therefore exhibit higher static power. Likewise, there will be devices that draw lower leakage current and therefore exhibit lower static power. The higher performance devices are generally also the ones that draw higher leakage current, and vice versa; this relationship is shown in [Figure 12](#).

Figure 12. Distribution of Leakage Power Characteristics and Static Power Consumption Across Semiconductor Devices, and Relationship to Performance



Normally, a semiconductor manufacturer like Altera might specify a maximum static power consumption for a given performance level. For example, consider a performance level that corresponds to the green area shown in Figure 12. The devices in the distribution that are to the right of the green area also meet that performance level, but they also consume higher static power. The standard practice would be to offer all devices that meet the performance level with their worst case static power determined by the devices to the right of the green area. However, Altera offers the devices in the green part of the distribution with a separate ordering code, indicated with a letter “L” to signify lower static power. The “L” devices draw up to 30 percent lower total static power than standard devices. The standard devices, which include all of the devices in the distribution that meet the performance level, are also offered for less power-sensitive customers; the standard devices are indicated with a letter “S”.

V_{CC} PowerManager

Arria 10 devices offer another option to tune device voltage, called V_{CC} PowerManager. With this feature, users can trade off performance to achieve even greater power reductions than are possible with SmartVoltage ID. Arria 10 devices enabled with the V_{CC} PowerManager can operate at either the standard V_{CC} voltage of 0.9 V or a set lower voltage level. The V_{CC} PowerManager devices also exhibit lower static power, similar to the “L” devices described before. When powered at the standard V_{CC} voltage level, these devices operate at the -2 speed grade, consume standard amounts of dynamic power, and up to 30 percent lower static power. At the lower voltage level, the devices consume correspondingly lower amounts of dynamic and static power, and operate at a performance level similar to the -4 speed grade. The V_{CC} PowerManager devices are labeled with “-2M” to indicate that they can operate at two voltages, and depending on the design, using this feature can reduce total power by up to 40 percent compared to designs that do not use the feature. The V_{CC} PowerManager is an effective technique for designs with high percentage of logic resource utilization, but do not require the highest performance logic fabric.

Table 3 summarizes the power reduction options offered with Arria 10 FPGAs and SoCs, as well as their speed grade and temperature grade support, device labeling, and power benefit.

Table 3. Power Reduction Options in Arria 10 FPGAs and SoCs

Feature	Speed Grade Support	Temperature Grades Supported	Power Benefit (vs. Standard Devices)
Programmable Power Technology	All	All	Up to 20% lower core static power
SmartVoltage ID	-3, -4	Industrial, Extended	Up to 40% lower core static power
Lower Static Power Devices	-2, -3	Industrial, Extended	Up to 30% lower static power
Standard Static Power Devices	-2, -3, -4	Industrial, Extended	None
V _{CC} PowerManager	-2	Industrial, Extended	Up to 30% lower static power at standard voltage <i>or</i> Up to 40% lower total power at lower voltage

Table 4 summarizes the power and performance benefit of Arria 10 devices compared to prior-generation devices.

Table 4. Power and Performance Benefit of Arria 10 Devices Compared to Prior-Generation Devices

Generation 10 Production	Comparison	Performance Increase	Power Reduction
Arria 10 FPGAs and SoCs	Arria V GX/GT FPGAs	60% faster	Up to 40% lower
	Stratix V FPGAs	15% faster	Up to 60% lower

Power Reduction in Stratix 10 FPGAs and SoCs

Stratix 10 FPGAs and SoCs will deliver significantly lower power than any other high-performance FPGA owing to the 14 nm Tri-gate process and a power-efficient architecture. Figure 13 shows the power reductions that Stratix 10 devices will demonstrate, relative to the prior-generation Stratix V FPGAs. Designs migrating from Stratix V standard devices can experience up to 55 percent reduction in total power with the application of any power reduction techniques. With the addition of power reduction techniques, power savings of up to 70 percent are possible.

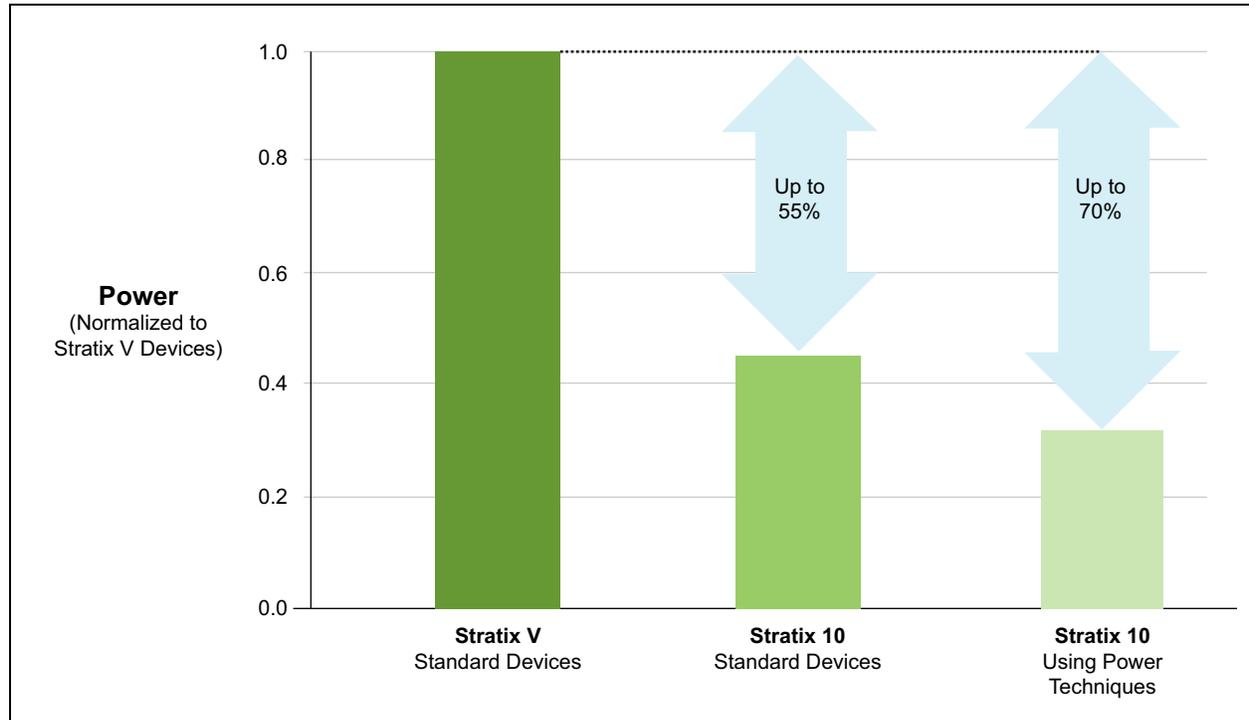
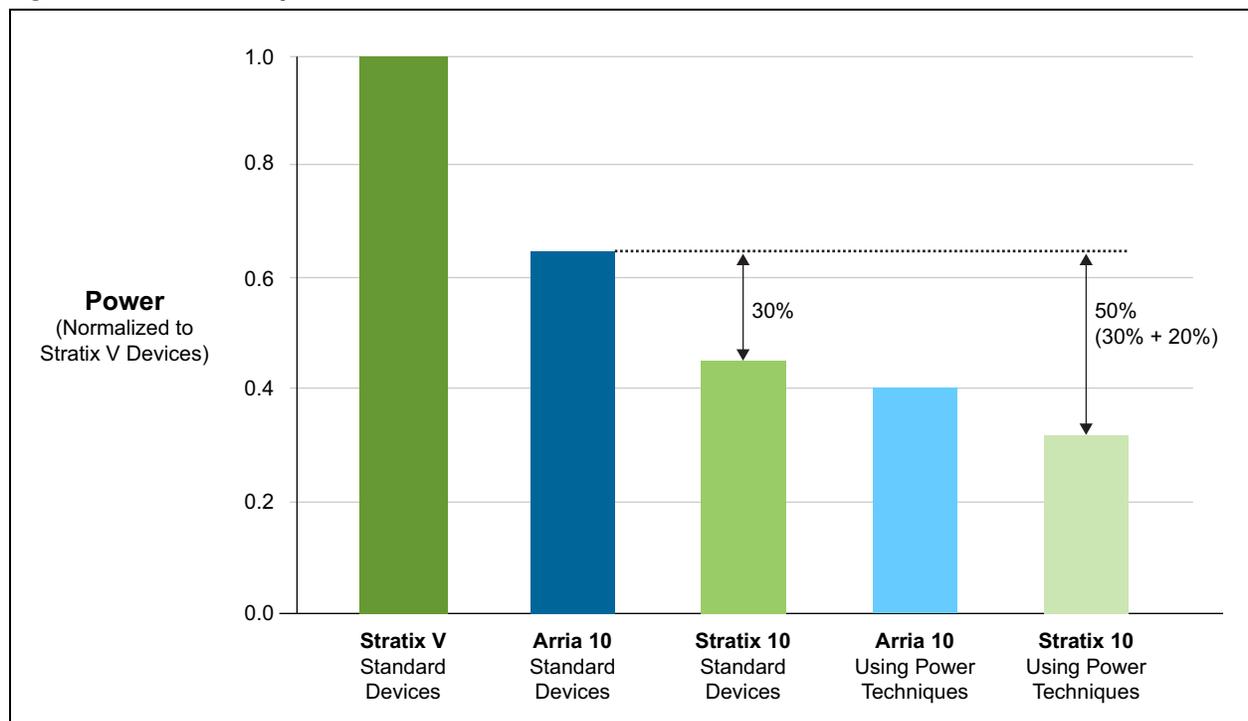
Figure 13. Power Reduction in Stratix 10 Devices Compared to Stratix V FPGAs

Figure 14 shows the power reductions that Stratix 10 devices will deliver relative to Arria 10 devices. The dark green Stratix 10 device bar and the light blue Arria 10 device bar shows that power reductions on the order to 40-50 percent from Stratix V device levels are possible from both Arria 10 devices and Stratix 10 devices. However, for Arria 10 devices, this level of reduction requires applying power reduction techniques to their maximum effect, whereas with Stratix 10 devices, this power reduction level derives simply from the process and architecture characteristics of Stratix 10 devices. With the application of power techniques, further reductions are possible such that Stratix 10 devices can reduce power up to 70 percent compared to the same design in a Stratix V device.

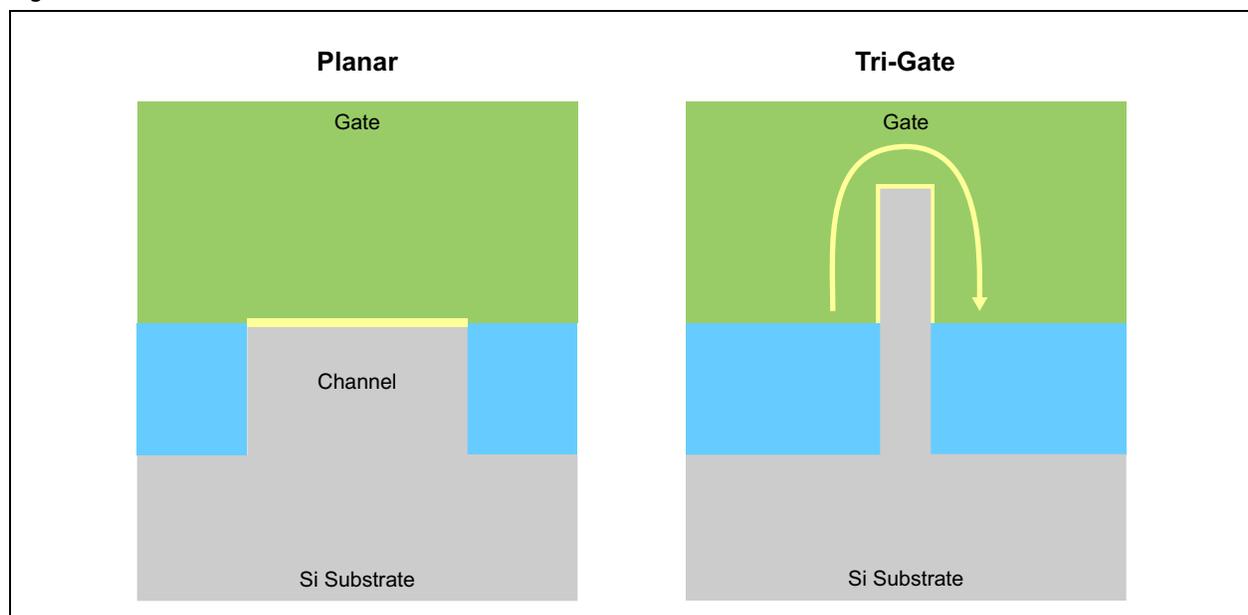
Figure 14. Power Consumption of Stratix 10 Devices Relative to Arria 10 and Prior-Generation Stratix V FPGA



Tri-Gate Process Reduces Static and Dynamic Power

Besides much higher performance, the Tri-Gate process also delivers major benefits in terms of lower power, both static and dynamic. The Tri-Gate transistor addresses the challenge of increasing leakage power by minimizing the leakage associated with the channel, as shown in Figure 15.

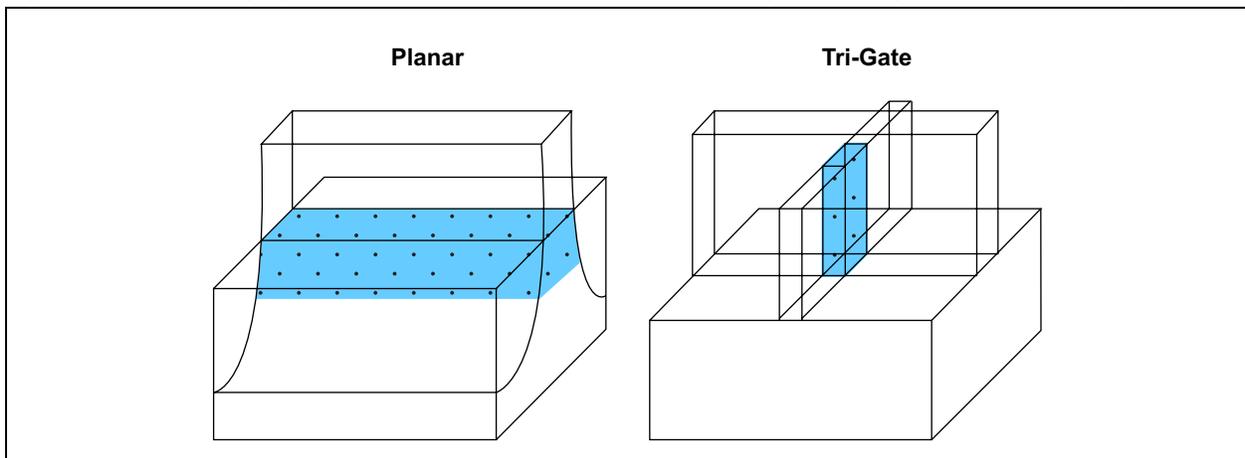
Figure 15. Tri-Gate Transistors Have More Effective Control over Channel Current



In the planar transistor, shown on the left in [Figure 15](#), the gate only influences the channel along a single surface. This single-surface influence limits the effectiveness with which voltage applied to the gate can shut off the current flowing in the channel. Undesired current flowing in the channel is leakage current, contributing to static power consumption. In the Tri-Gate transistor, shown on the right, the gate surrounds the channel on three sides, and is able to control the flow of current through the channel much more effectively.

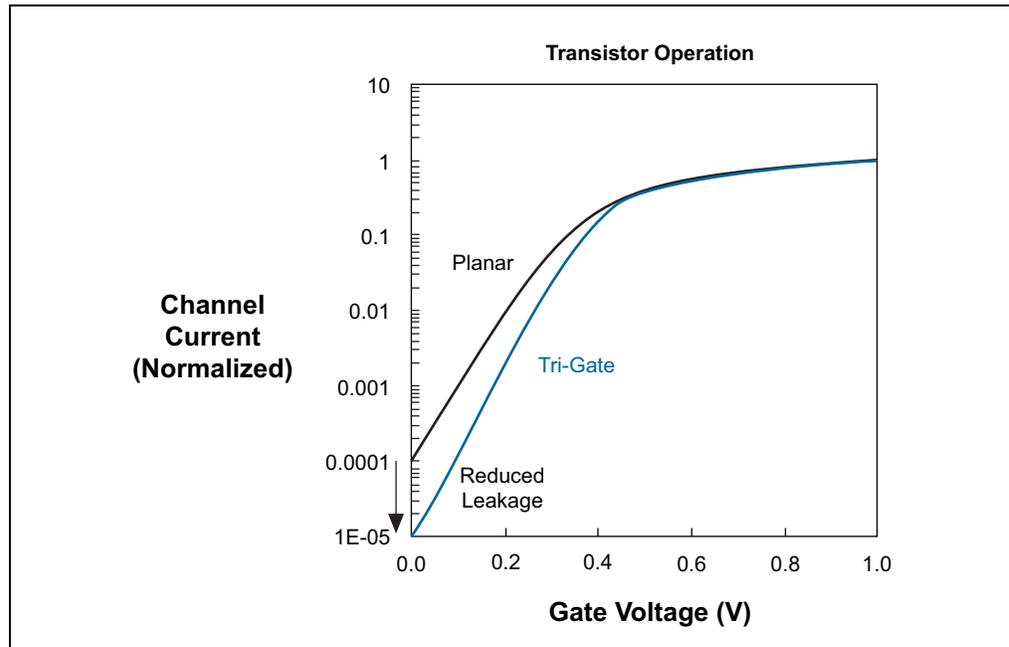
[Figure 16](#) shows another way that the Tri-Gate transistors reduce leakage current. In a planar transistor, shown on the left side in [Figure 16](#), the channel is doped to set the threshold voltage and to limit the flow of leakage current through the channel. At submicron levels, even small fluctuations in doping can result in undesired variation in leakage current. The Tri-gate transistors require little or no channel doping to control leakage since the gate influences the channel from both sides of the channel (fin). The reduced doping in the Tri-gate channel decreases variability resulting in lower worst-case and average leakage current.

Figure 16. Reduced Doping in Tri-Gate Transistors vs. Planar Transistors (Intel Developer Forum, June 2012)



The effectiveness of the Tri-Gate transistor in minimizing leakage power is shown in [Figure 17](#), which illustrates the channel current in a Tri-Gate transistor vs. a planar transistor as a function of gate voltage. As shown on the far left side of the graph, when the gate voltage is at 0 V, there is an order of magnitude lower amount of leakage current flowing through the channel of the Tri-Gate transistor compared to the planar transistor.

Figure 17. Current in Planar and Tri-Gate Transistors as a Function of Gate Voltage (Intel Developer Forum, June 2012)



Tri-Gate transistors exhibit lower active or dynamic power relative to planar transistors as well, due to lower supply voltages required. Figure 18 shows this active power reduction; the trend in active power across process nodes has been in the downward direction. However, as demonstrated by the bend in the curve further downward from the 32 nm planar node, the introduction of Tri-Gate transistors has clearly further reduced the dynamic power beyond the trend established by prior process geometry shrinks.

Figure 18. Active Power in Planar and Tri-Gate Transistors (Intel Developer Forum, June 2012)

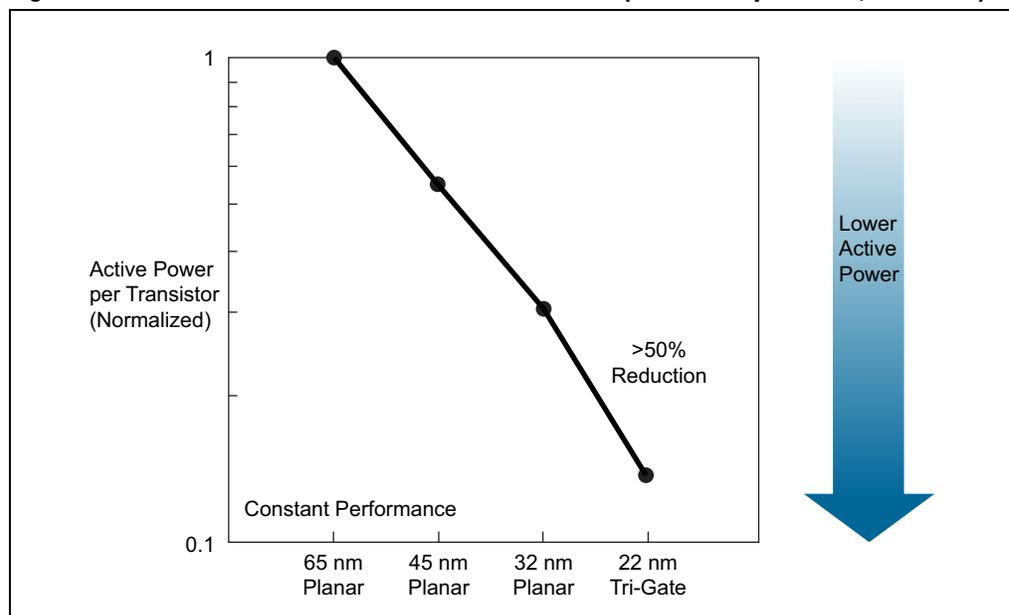


Table 5 summarizes the performance and power benefits of Stratix 10 devices compared to prior-generation high-end FPGAs (Stratix V FPGAs).

Table 5. Performance and Power Benefits of Stratix 10 Devices Compared to Prior-Generation FPGAs

Generation 10 Products	Comparison	Performance Increase	Power Reduction
Stratix 10 FPGAs and SoCs	Stratix V FPGAs	Over 2x faster	Up to 70% lower

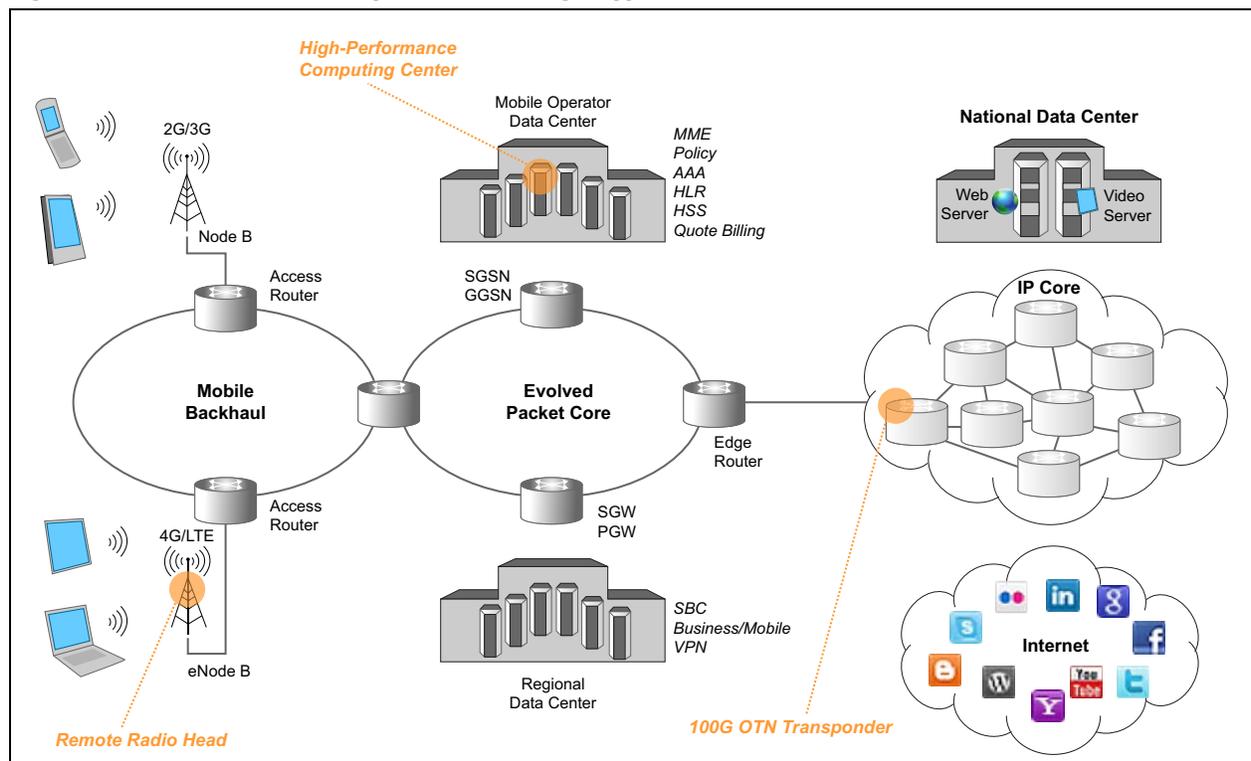
Applying the Power and Performance of Next-Generation FPGAs and SoCs to ICT Challenges

To evaluate the impact of the performance and power capabilities of next-generation FPGAs and SoCs, consider three application areas in an ICT infrastructure example, shown in Figure 19.

The three application areas are:

- High-performance computing in a data center
- 100G Optical Transport Network (OTN) transponder in the network core
- Remote radio head

Figure 19. ICT Infrastructure Example with Three Target Application Areas



High-Performance Computing in a Data Center

In a data center, raw processing power is highly attractive, in particular the highest processing power per watt, so high power efficiencies are extremely desirable. To understand the potential for FPGA-based power or performance optimization in data centers, we can analyze the example of a Smith-Waterman algorithm. The Smith-Waterman algorithm is commonly used for bioinformatics applications such as genomic searches, which are extremely compute and memory intensive. A typical approach to applying Smith-Waterman in a research or commercial application is to use general-purpose processors or GPUs, and results are reported in Cell Updates per Second (CUPS). Using Open Computing Language (OpenCL™)—an open programming model for accelerating algorithms across multiple platforms like CPUs, GPUs, DSPs, and FPGAs—a highly optimized FPGA-based approach to Smith-Waterman can deliver power efficiencies far beyond what is possible with CPUs and GPUs ⁽¹⁵⁾. Table 6 lists the actual tested results of this FPGA-based approach using OpenCL, in comparison to more common implementations.



OpenCL and the OpenCL logo are trademarks of Apple Inc., and used by permission by Khronos.

Table 6. Three Different Implementations of Smith-Waterman Algorithm and Their Associated Power Efficiencies; (Test, Sample) = (256, 15M) Sequences

Platform	Throughput (MCUPS)	Power (Watts)	Efficiency (MCUPS/Watt)
Intel® Xeon® Quad-	40	140	0.29
NVIDIA GT620	438	50	8.76
Stratix V A7 FPGA	32,596	25	1,303

Table 7 shows that using today's highest speed FPGA technology, a Stratix V device, it is possible to achieve an implementation that is over 148 times more power efficient than the GPU-based approach.

Table 7. Smith-Waterman Algorithm Implemented in Arria 10 and Stratix 10 Devices, Expected Throughput, Power Consumption, and Power Efficiencies; (Test, Sample) = (256, 15M) Sequences

Platform	Throughput (MCUPS)	Power (Watts)	Efficiency (MCUPS/Watt)
Arria 10	>35,000	18	>1,900
Stratix 10	>70,000	12	>5,800

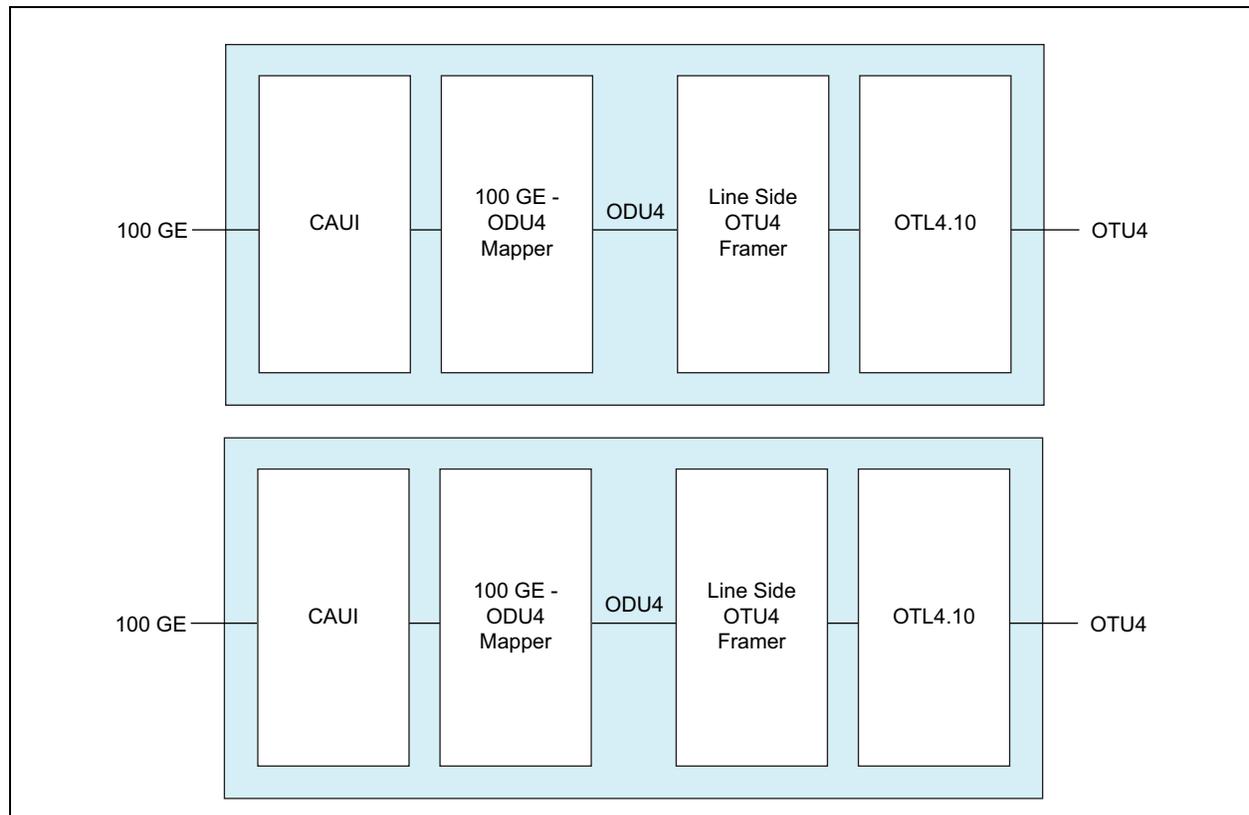
Table 7 shows the estimated results of Smith-Waterman running in both Arria 10 and Stratix 10 FPGAs. The same design running in an Arria 10 FPGA would draw less than 18 Watts, and the resulting power efficiency would be over 200 times better than the GPU-based approach. In Stratix 10 FPGAs, the power efficiency is over 660 times better than the GPU-based approach. This kind of improvement is an indicator of the kinds of power efficiencies that data center operators can achieve when applying the high customizability and performance of FPGAs and SoCs to the most compute-intensive functions that their servers handle, and the OpenCL programming model means that these operators can simply take a familiar C-based approach to implement their algorithms for FPGAs or any of the typical target platforms like CPUs, GPUs, or DSPs.

100G OTN Transponder

Moving into the network core, a wide variety of OTN functions can benefit from FPGA power or performance optimizations. A specific case is the OTN transponder, which converts signals from one wavelength to another for optical transport.

Consider an OTN transponder that converts a 100 GbE signal to the appropriate line rate for optical transport (OTU4). A single high-end FPGA such as a Stratix V A5 device can implement two of these transponders in a single chip, shown in Figure 20.

Figure 20. Implementation of Two 100G OTN Transponders in a Single Stratix V FPGA or Arria 10 FPGA

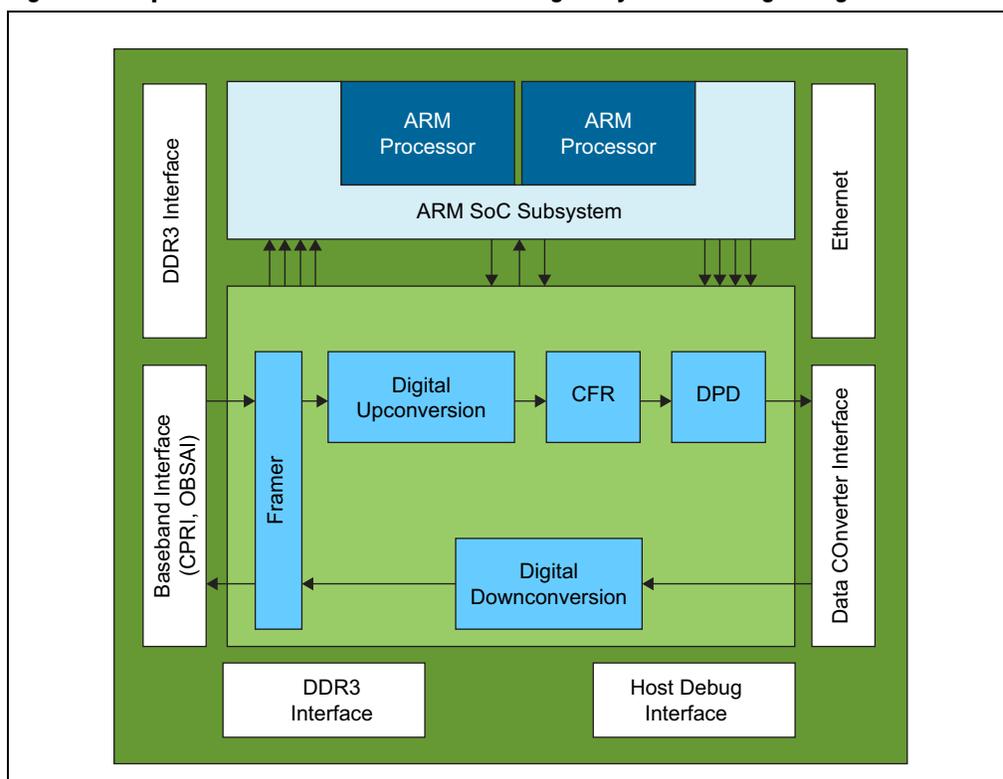


The power consumption of this single-chip solution of two 100G OTN transponders is 31.4 W in a standard Stratix V A5 device. The same design implemented in a single Arria 10 (10AX057) device enabled with the V_{CC} PowerManager feature consumes less than 19 W, a reduction in total power of 40 percent. Furthermore, in Stratix 10 FPGAs, the same functionality consumes less than 14 W, a savings of 54 percent compared to the current state-of-the-art FPGA.

Remote Radio Head

Remote radio heads (RRHs) are found at the edge of the network, transmitting data to and receiving data from mobile devices. RRHs have an additional design challenge in their thermal environment and compact size. Situated in remote locations as their name indicates, exposed to outside elements and harsh conditions, radio heads must dissipate their heat with little or no airflow, which in most situations results in a practical power limit of 20 W or less. FPGAs and programmable SoCs are optimal for this application owing to their integration capabilities, enabling most or all digital functions in the radio head to be consolidated into a single chip, as shown in Figure 21.

Figure 21. Implementation of Remote Radio Head Digital System in a Single Programmable SoC



For the coming generation of radio heads, a clock frequency of nearly 500 MHz is desirable to provide the oversampling required to maximize radio frequency (RF) bandwidth. In today's FPGA technology, a representative multistandard radio head in a two transmit and two receive configuration (2T2R) with a 60 MHz bandwidth may consume well over 20 W, exceeding the strict limits imposed by the RRH thermal conditions. In an Arria 10 device, however, this type of RRH can consume less than 18 W, while still maintaining the maximum level of integration of all digital functionality into a single chip.

For Stratix 10 devices, the smallest device available would provide more resources than needed for this application, so it would be a better fit for RRHs offering higher bandwidth on the order of 100 MHz or integrating larger antenna configurations (4T4R, 8T8R). For these RRHs, targeted clock frequencies starting at 736 MHz are also well within Stratix 10 FPGA and SoC capabilities, making them a good fit for satisfying the demanding performance requirements while meeting the 20 W power maximum limit.

Conclusion

Today's ICT equipment developers face a daunting problem in addressing exponential growth in bandwidth demand while minimizing power consumption. Fortunately, the most advanced FPGAs and SoCs of the coming generation are designed to help system designers tackle this challenge. By leveraging multiple process technologies and revolutionary approaches to transistor design, as well as new architectures and comprehensive device-level power features, these devices enable customization of products at the component level to achieve the optimal balance of power and performance. Table 8 summarizes the Generation 10 product performance or power benefits in the context of the ICT application areas discussed; these are just a few of the places in the entire ICT technology sector that can benefit from next-generation FPGAs and SoCs.

Table 8. Summary of Performance or Power Benefits Provided by Next-Generation FPGAs and SoCs in an ICT Application Area Example

Generation 10 Products	Application Area	Function	Power or Performance Benefit
Arria 10 Devices	Data center	High-performance computing	>148x higher power efficiency vs GPUs
	Core network	100G OTN transponder	40% power reduction
	Mobile communications	60 MHz remote radio head	500 MHz within 20 W thermal budget
Stratix 10 Devices	Data center	High-performance computing	>200x higher power efficiency vs GPUs
	Core network	100G OTN transponder	65% power reduction
	Mobile communications	60 MHz remote radio head	736+ MHz within 20 W thermal budget

While it seems a certainty that the world is headed towards explosive growth in bandwidth, the path to the Zettabyte era may not always be clear. Power and performance are just some of the most visible challenges, but the most dangerous are those yet to be identified. Responding to known challenges is a minimum expectation for any enterprise, but the ability to respond to unforeseen shifts and disruption can mean the difference between success and failure—this may be where the flexibility of execution and implementation afforded by programmable logic provides the greatest benefit.

References

1. International Telecommunications Union, The World in 2011 Facts and Figures: www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2011.pdf
2. Telegeography Global Bandwidth Research Service: www.telegeography.com/products/commsupdate/articles/2012/07/18/international-bandwidth-demand-grows-45/

3. International Telecommunications Union, The World in 2013 Facts and Figures:
www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf
4. Cisco Visual Networking Index
www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html
5. 2012 Annual Report of the Centre for Energy-Efficient Telecommunications (CEET):
www.ceet.unimelb.edu.au/pdfs/ceet_annualreport_2012.pdf
6. Fujitsu ICT Sustainability: The Global Benchmark 2012 report:
<https://www-s.fujitsu.com/global/solutions/sustainability/Fujitsu-Sustainability.html>
7. Computer Weekly:
www.computerweekly.com/news/2240164589/Datacentre-power-demand-grew-63-in-2012-Global-datacentre-census
8. Sustainable ICT in Corporate Organizations, International Telecommunications Union:
www.itu.int/dms_pub/itu-t/oth/4B/04/T4B0400000B0011PDFE.pdf
9. Smart 2020 Report:
www.smart2020.org/_assets/files/01_Smart2020ReportSummary.pdf
10. Google; Data Centers Efficiency:
www.google.com/about/datacenters/efficiency/internal/index.html#servers
11. Growth in Data Center Electricity Use 2005 to 2010, Koomey:
www.koomey.com/post/8323374335
12. Reducing the Carbon Footprint of ICT Devices, Platforms and Networks; Thierry Van Landegem, Chair of the GreenTouch Operations Committee:
www.greentouch.org/uploads/documents/Van%20Landegem%20GeSI%20Reducing%20Carbon%20Footprint%20May%202012.pdf
13. Intel Reinvents Transistors Using New 3-D Structure:
newsroom.intel.com/community/intel_newsroom/blog/2011/05/04/intel-reinvents-transistors-using-new-3-d-structure
14. ITRS 2009 Executive Summary:
www.itrs.net/Links/2009ITRS/2009Chapters_2009Tables/2009_ExecSum.pdf
15. OpenCL for Altera® FPGAs:
www.altera.com/products/software/opencl/opencl-index.html

Acknowledgements

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

Document Revision History

Table 9 shows the revision history for this document.

Table 9. Document Revision History

Date	Version	Changes
June 2013	1.0	Initial release.