XILINX

WP528 (v1.0)

Achieving Embedded Design Simplicity with Kria SOMs

Xilinx Kria SOMs' unique accelerated application-based approach offers a new paradigm for software-based design while maintaining system-level flexibility and FPGA performance benefits for applications in Industrial, Vision, and Healthcare & Sciences markets.

ABSTRACT

Kria™ System-on-Modules (SOMs) are an integrated embedded, pluggable, credit-card-sized, Arm® SoC-based solution that can readily support a variety of pre-defined and user-defined applications. Kria SOMs and the connected Xilinx App Store are intended to be a "welcome" sign for embedded and AI software developers that have wanted to adopt the benefits of FPGA technologies to create innovative ideas in production systems but do not know, or even want to know, where to begin. Additionally, for the traditional Xilinx designer, Kria SOMs are a convenient and efficient hardware platform to deploy production designs.

Introducing Kria SOMs

FPGA capabilities continue to expand year after year, building upon a core value proposition of flexibility and performance advantages over other embedded technologies. In recent years, unquestionably, the most popular trend in FPGAs is the integration of embedded processors and peripherals, enabling designers to assign the right tasks to the right computing engines. The resulting adaptive SoCs offer three degrees of freedom to users: software programmability, hardware programmability, and scalability for embedded platforms by offering a consistent processing sub-system across varying amounts of FPGA fabric. For these reasons, among others expanded in Xilinx white paper WP493, there has been widespread adoption of Zynq®-7000 SoCs and Zynq UltraScale+™ MPSoCs in Industrial & Healthcare IoT production systems over the past decade.

This adoption is due to the clear and differentiating benefits of FPGA technology. However, as exciting as this has been, much broader adoption is possible. The two most common blocking factors to wider adoption are 1) unfamiliarity with traditional FPGA design flows, and 2) cost of the iterative FPGA design process measured in terms of development time. There is a strong indication that many system architects and embedded and AI software developers have heard about the benefits of FPGAs and want to access them—but without the difficult and lengthy process of learning 'new-to-them' design flows, and focusing on embedded hardware aspects of the development process. Even when it comes to FPGA experts, many have indicated a desire for an easy and cost-effective production hardware deployment option for their embedded designs. This desire has grown even stronger due to the collaborative effort required to bring up embedded boards in co-located electronics labs and the corresponding limitations of being able to do so in the wake of a global pandemic in the years 2020 and 2021.

Kria SOMs are the latest addition to a ready-for-production portfolio of Xilinx products that already consists of accelerator cards for x86 applications, FPGAs, and Adaptive SoCs. Kria SOMs are an integrated embedded, pluggable, credit-card-sized, Arm SoC-based solution. The premise of the Xilinx SOM architecture is to streamline inclusion to only the SoC, memories, power supplies, and other fundamental support for the SoC. The integration into a production-deployed system happens when the SOM is mated with a low-complexity, end-user designed, printed circuit board carrier card, leveraging a simple set of open connectors, that addresses the connections and additional components specific to that user's end system. For development and evaluation purposes, Xilinx offers a starter kit that consists of a non-production Kria SOM mated to a pre-defined carrier card with various interfacing options to connect sensors or other peripherals. The starter kit is the easiest way to get started with Kria SOMs, especially with Vision AI applications, and provides a path to move to a custom user-defined production carrier card that is developed for the target application.

Save up to Nine Months of Development Time and the Associated Cost

Kria SOMs offer a uniquely simplified user experience for those new to FPGA and SoC technology and provide compelling efficiency gains to long-time FPGA users seeking to buy, rather than make, their production hardware. For the FPGA expert, Kria SOMs save up to nine months of development time and the associated cost—a very straightforward value proposition. These nine months saved are typically spent in device selection, schematic entry, board layout, multi-disciplinary board bring-up, and sometimes partial re-design. The more interesting and much greater gains await the users who are new to Adaptive Computing; accordingly, the balance of the white paper focuses on those gains.

Kria SOMs represent a new mainstream product line for Xilinx, with the launch of the first member SOM (the K26) and the first starter kit (the KV260 Vision AI Starter Kit). See Figure 1.
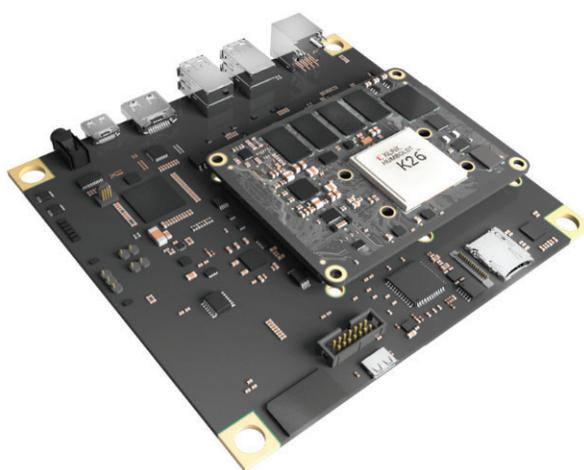
WP528_1_030421

*Figure 1:* **Xilinx K26 Kria SOM and KV260 Vision AI Starter Kit**

Developers can expect more Kria SOMs and kits to be added over time, and while this white paper focuses on the first offerings, the same concepts apply to future offerings. For product-specific information on Kria SOM offerings, go to: https://www.xilinx.com/products/som/kria.html.

A key enabler in achieving embedded design simplicity with Xilinx Kria SOMs lies in the application-specific infrastructure that is provided with each starter kit, which can be carried over into the user's production design. The first Kria starter kit, the Kria KV260 Vision AI Starter Kit focuses on Vision AI applications. Future offerings focus on a wide array of other applications, both developed by Xilinx and by third parties, and are available from the Xilinx App Store. The Xilinx App Store delivers a production-grade set of reference designs called Kria accelerated applications. This unique accelerated application-based approach offers a new paradigm for software-based design while maintaining FPGA performance benefits and system-level flexibility for targeted applications.

Simplifying FPGA Design Using the Accelerated Applications Approach

The double-edged sword of Adaptive SoCs is that they are more flexible than any other mainstream semiconductor technology available but can add complexity for the new user. This flexibility manifests itself in many ways: programmable logic, programmable software, programmable I/O, programmable DSP Engines, programmable embedded RAM, programmable security, programmable power consumption, etc. Kria SOMs, by design, reduce the burden of this flexibility, focusing on the most common and mainstream use cases. These use cases have been developed into accelerated applications and can ease the user into the design community via the Xilinx App Store. The transformative impact of the accelerated applications is similar to how the invention of the automatic transmission made driving an automobile much more accessible than a manual transmission. There will always be manual transmission enthusiasts, but most of the world has naturally migrated to what is easiest to operate and most convenient. Even with this streamlining of the inherent capabilities of a Xilinx device, there is still significantly more system-level flexibility available with Kria SOMs than either microprocessor or GPU-based SOMs available in the market.

Beyond the key convenience features of the pre-configured power supply, the automated boot process with "golden" image fallback and I/O assignments, Kria SOMs provide a wealth of additional features to accelerate development. There is support for a pre-built and scalable Linux infrastructure, whether the user's preference is the Ubuntu and Ubuntu Core distributions by Canonical or Xilinx-managed Linux kernel with PetaLinux tools, and capability for custom builds using Yocto. Both come with a set of very practical system utilities such as temperature, performance monitors, quality-of-service configuration for DDR memory and other peripherals, over-the-air firmware update capabilities, and built-in-self-test, to name a few. These features greatly contribute to the simplicity of using Kria SOMs. However, the most significant attribute of Kria SOMs is the growing library of accelerated applications that are found on the Xilinx App Store, both for free and for a fee, and can easily be loaded into a Kria SOM. See Figure 2.
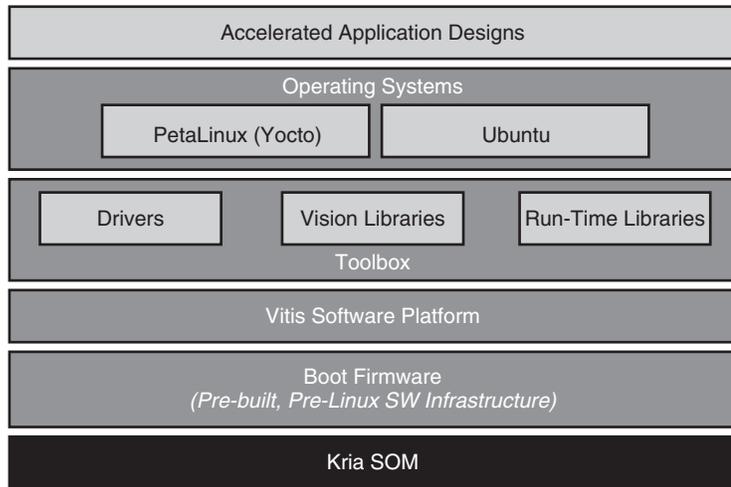
WP528_2_041521

Figure 2: **Pre-Built Linux Infrastructure Simplifies Embedded Design**

Pre-Built Programmable Logic and Accelerated Applications

Kria accelerated applications are intended to create robust application-specific FPGA hardware designs with a basic software application to be extended by the end user. In an accelerated application, the programmable logic portion of the SoC is pre-built for the user. No modification is required to use it, although it is certainly possible if desired.

Xilinx has created a number of accelerated applications that can be modified for the designers' diverse requirements. The initial Kria accelerated applications are based on research that identified the most popular Vision AI applications that users have been developing. These accelerated applications by Xilinx can be modified by customers based on their target needs, saving them software development time when compared to starting from scratch. More accelerated applications will be added to the Xilinx App Store, both by Xilinx and the ecosystem (who can deliver applications leveraging their expert knowledge in their respective focus areas). See Table 1.

Table 1: **Accelerated Applications for Kria SOMs**

Kria Accelerated Application	Description
Smart Camera	Supports Face Detection[1] + Network & Display
AI Box, ReID	Multi-Stream Face & Pedestrian Detection[1]
Distributed Smart Camera, AI Box	Distributed, Scalable, Multi-Stream Tracking and Re-Identification
Machine Vision Camera	OpenCV-Based Defect Detection
Natural Language Processing	Real-time Audio Keyword Spotting w/ Vision AI Fusion

Notes:
1. Supports Xilinx Model Zoo.

In the case of the first starter kit offering, Vision AI applications, such as security, smart cities, traffic management, and retail analytics, are the focus applications. These applications have driven design decisions on the KV260 Starter Kit, including elements like the chosen connectors for I/O interfacing. The connections include interfaces such as MIPI, HDMI, DisplayPort, USB, and Ethernet. Future starter kits will target any number of applications, such as machine vision, vision-guided robotics, and beyond. Kria accelerated applications enable users to rapidly move from the evaluation phase to design and then deployment.
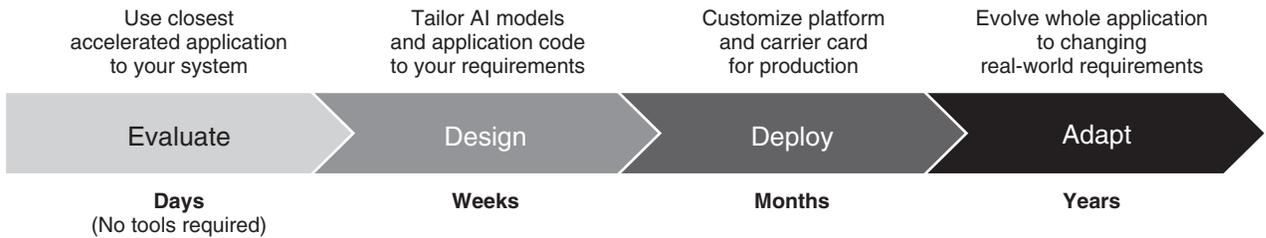
Typical Flow from Evaluation to Production Deployment

Here is an example of a typical flow from evaluation to production deployment:

1. First days (Evaluate): Choose the closest Xilinx pre-built design (accelerated application) that meets the user's requirements and is a good fit to assess the capabilities of the platform.

2. First weeks (Design): Customize the application software and the AI models to tune the Kria accelerated applications to the end product requirements. Users can also optimize the complete image datapath for the user's choice of AI model with Vitis™ libraries' hardware accelerators for maximizing performance.

3. First months (Deploy): Customize the hardware to user-specific peripherals, such as different camera sensors, and start the process of building a custom carrier card. As required, FPGA hardware can also be modified at this stage via the Vitis design tools using the ample online tutorials and design guides.

4. Years (Adapt): Update hardware and software in the field via over-the-air updates to maximize the relevancy of the product and potentially generate new revenue streams for the user's company as a result of these updates. Evolving standards, threats, and customer requirements can drive changes to AI models, vision pipeline, control loops, communication protocols, and encryption engines. Kria SOMs offer the flexibility and built-in capability to make these changes after deployment.

See Figure 3.

Use closest accelerated application to your system	Tailor AI models and application code to your requirements	Customize platform and carrier card for production	Evolve whole application to changing real-world requirements
Evaluate	Design	Deploy	Adapt
Days (No tools required)	**Weeks**	**Months**	**Years**

WP528_3_032921

Figure 3: **Kria SOM Flow - From Evaluation to Production**

As presented in the above scenario, no Xilinx tools need to be installed during the evaluate phase. Software tools, such as a GNU compiler and Vitis AI, are used for the second step, the design phase. The third phase is deployment. If hardware customization is required anywhere, it will most likely be in this phase—so FPGA tools, such as Vitis and/or Vivado® Design Suite, might be required.

Benefits of Kria Accelerated Applications

Kria accelerated applications are important because they enable users to start with familiar embedded software and AI tools in an Ubuntu or Yocto Linux environment instead of having to engage with potentially unfamiliar FPGA tools. For a user to take advantage of this convenient flow, the premise is that there is an accelerated application in the Xilinx App Store that matches the needs of the user (e.g., smart camera, AI box, machine vision camera with the KV26 and KV260—and robotic controller and electric drives in the future). Application software and AI processing are the primary modifications required for production. As a result, FPGA customization is not required but is fully supported.

With software design convenience, the application-specific FPGA acceleration blocks, pre-built in these systems, provide world-class performance and low-latency. The same SOM can dramatically change "personalities" in the form of acceleration blocks presented to the user depending on which accelerated application is loaded into the SOM. Furthermore, each accelerated application can be customized using Vitis accelerated libraries, ranging from OpenCV to a wide assortment of other common functions. This flexibility is the fundamental value proposition of the SOM's adaptive foundation.

Key Performance and Features of Kria SOMs

The key features and figures of merit vary depending on the Kria accelerated application used and the functionality associated with that application. In the context of the first offering, the K26 SOM and the KV260 Vision AI Starter Kit are focused on AI throughput, performance/watt, and cost/channel. A more detailed performance benchmarking report is available in the companion

white paper: WP529, *Kria K26 SOM: The Ideal Platform for Vision AI at the Edge*. This white paper offers an overview of the advantages over competing SOMs and highlights specific use cases, such as license plate/numeric plate recognition in smart city applications.

The K26 SOM offers approximately a 3X performance advantage over Nvidia Jetson Nano, in terms of raw trillions of operations per second (TOPS). It also offers up to 2X performance/watt advantage over Nvidia Jetson TX2. Its low latency and high-performance deep learning processing unit (DPU) provides a substantial advantage over Nano and TX2 with networks in the machine learning performance benchmarking suite. See Figure 4.
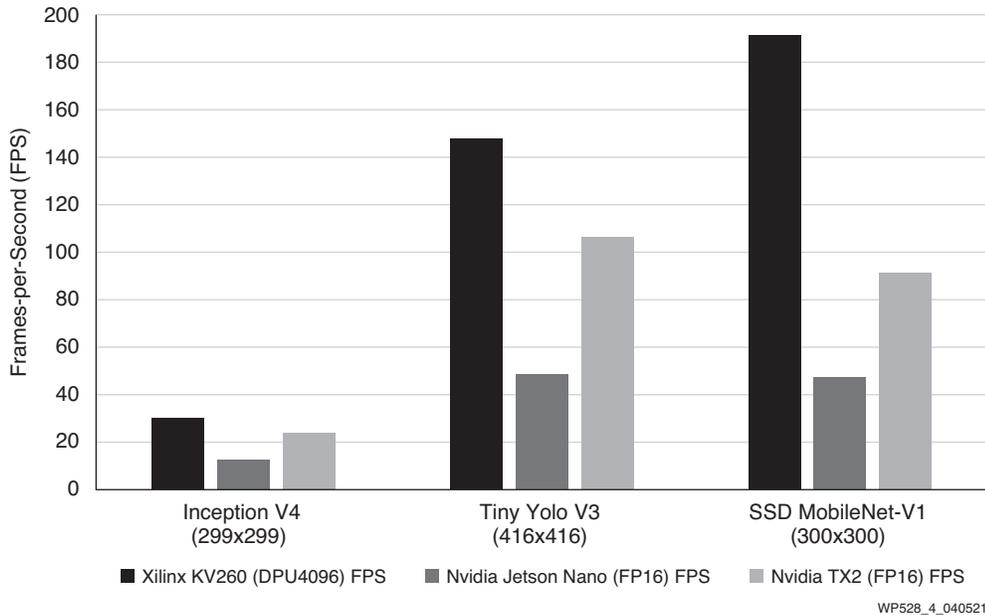

WP528_4_040521

Figure 4: **Xilinx Kria SOM vs. Nvidia Jetson AI Performance**

Kria K26 SOM Features

The K26 SOM leverages the Zynq UltraScale+ MPSoC architecture. As such, it and all K2x-series SOMs boast a 64-bit quad-core Arm® Cortex™-A53 application processor complex with a 32-bit dual-core Arm Cortex-R5F real-time processor and an Arm Mali™-400MP2 3D graphics processor. Included on the SOM is 4GB of 64-bit DDR4 memory along with QSPI and eMMC memory. The inherent secure boot capabilities with built-in hardware root-of-trust of the Zynq UltraScale+ architecture is extended via an external TPM2.0 for measured boot and is IEC 62443 capable (see WP513, *Xilinx IEC 62443 Compliant Product Enablement*). Furthermore, the K26 SOM can support up to a 4096-sized DPU with its 256K System Logic Cells in the FPGA fabric and has a built-in Video Codec for H.264/H.265. The K26 SOM supports a plentiful amount of 1.8V, 3.3V single-ended, and differential I/Os with a quad of 6Gb/s transceivers, and a quad of 12.5Gb/s transceivers. Together, there is support for a wide assortment of standards, such as MIPI, SLVS-EC, sub-LVDS, DisplayPort, HDMI, PCIe®, USB2.0/3.0, and much more, including user-defined standards. The size of the K26 SOM is 77mm x 60mm x 11mm, supporting ruggedized applications with future SOMs planning aggressive size reductions. The K26 SOM supports a commercial temperature rating of 0°C to +85°C junction temperature, as reported by the internal temperature sensor to the application processor, with all other devices on the SOM falling in line accordingly to that single measurement. Similarly, there is support for –40°C to +100°C in the industrial grade version. The SOM includes two 240-pin connectors to the carrier card. See Table 2.

Table 2: **K26 SOM Overview**

COMPUTE	
Application Processor	64-bit Quad-Core Arm Cortex-A53
Real-Time Processor	32-bit Dual-Core Arm Cortex-R5F
Graphics Processor	Arm Mali-400MP2
Programmable Logic	256K System Logic Cells
Deep Learning Processor	4K INT8 (upgradeable to INT4)
Video Codec (H.264/H.265)	Up to 32 Streams (total resolution ≤ 4Kp60)
Memory	26.6Mb On-Chip SRAM
Security	IEC62443 Security w/HW Root-of-Trust
INTERFACES	
Camera	11 x4 Full MIPI or sub-LVDS Interfaces 1 x4 SLVS-EC Interfaces
USB	4x USB 2.0 / 3.0
Multi-Media	DisplayPort, HDMI
Network	1Gb up to 40Gb Ethernet (w/GigE Vision)
Memory Interface	4GB 64-bit DDR4
Transceivers	4x 12.5Gb/s, 4x 6Gb/s
Mechanical	77 x 60 x 11mm w/ dual 240-pin connectors

KV260 Starter Kit Features

To streamline cost and simplicity of evaluation, while there are two connectors on the production K26 SOM, only one of those is populated for use with the feature limited SOM that is included in the KV260 Starter Kit. Other differences between the production SOM and the starter kit are the removal of eMMC memory in lieu of an SD card interface on the carrier card and the room-temperature-only operation of the starter kit.

The KV260 Vision AI Starter Kit supports a trio of MIPI sensor interfaces, one Imager Access System (IAS) connector from On Semiconductor with an onboard image signal processor (ISP), one additional IAS connector, and one Raspberry Pi connector. Multiple USB 2.0/3.0 interfaces are present along with Ethernet, HDMI, DisplayPort, and Pmod connectors to evaluate a wide assortment of Vision AI applications. The convenience of the starter kit enables unboxing to application development in well under one hour. See Figure 5.
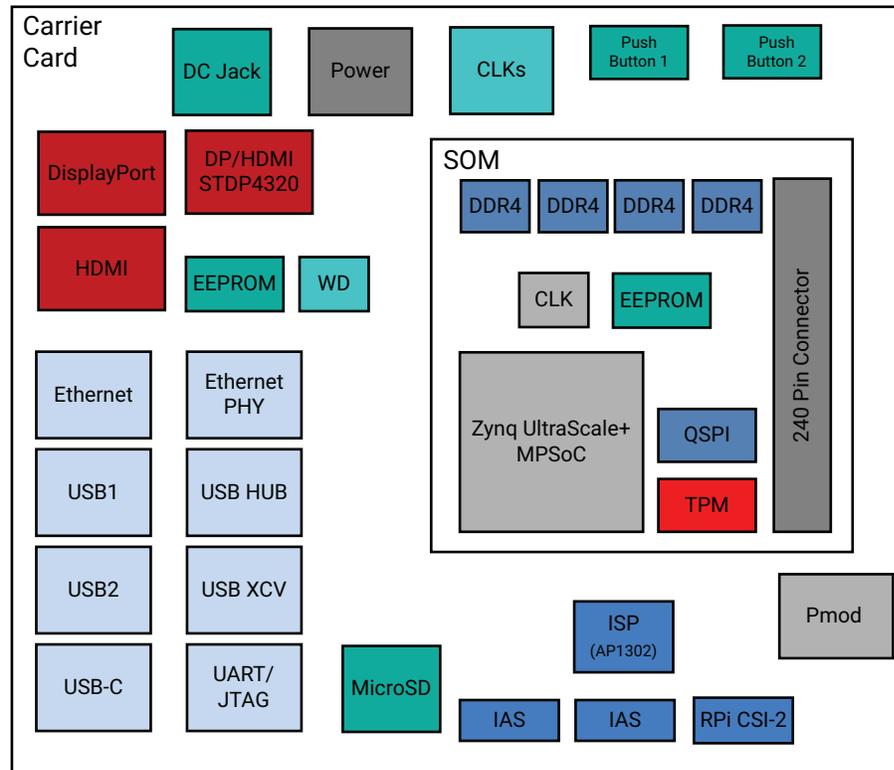
X24612-040621

Figure 5: **Block Diagram of KV260 Vision AI Starter Kit**

System-level Flexibility for Futureproof Products with Kria SOMs

In addition to the application-specific solution set and the intrinsic performance features that make Kria SOMs a compelling choice for new embedded designs, there are also several long-term business and technology benefits for the Industrial market.

The Industrial market is characterized by a requirement for long life cycles in harsh environments. Long life cycles, in this context, refer to long operating life, over extended temperatures, humidity, shock, and vibration. An Industrial grade Kria SOM is built to withstand 10 years of operation at 100°C junction, at 80% relative humidity. It can also withstand up to 40g of shock and 5g RMS of vibration.

To some, long life cycles also refer to availability. In the case of Xilinx SOM products, availability is a minimum of 10 years in the Industrial grade offering. This is a key advantage to Kria SOMs as a variety of components, including the DDR4, QSPI, eMMC, TPM2.0, and power management devices can be volatile in terms of both price and availability. Xilinx offloads the need for the user to pay attention to these details because a long life cycle is one of the many values of Kria SOMs. The Kria SOM model intrinsically offers customers the ability to transition from SOM to chip-down as business dictates due to the rise or fall of shipments.

From a technology perspective, Kria SOMs with their adaptive device foundation, offer the user significant flexibility advantages. This system-level flexibility can be further enhanced through future updates to accelerated applications. For example, an extrapolation of DPU performance

shows that moving from INT8 to INT4 offers up to a 77% performance boost on hardware with reduced logic and lower on-chip memory utilization. Reduced precision inference architectures, such as Fast, Scalable Quantized Neural Network Inference (FINN), can be employed as well as support the notion that Kria SOM hardware gets better over time and can be upgraded after deployment in the field. This is in contrast to GPU architectures, which are fixed.

Adaptable to Market Trends

Sometimes the market drives the migration from one technology or interface to another, and Kria SOMs can adapt to these external forces as well. Xilinx's programmable I/Os and internal programmable logic can support the transition from an imaging standard like MIPI to SLVS-EC or even from SLVS-EC Version 1.0 to Version 2.0, which supports twice the data rate.

Another element of the Kria SOMs' flexibility is that they are cloud-native. They offer users the flexibility to distribute workloads across the edge and cloud rapidly and seamlessly. Other IoT solutions end up being too constrained to reliably be counted on for handling cloud workloads when faster response times are desirable. Kria SOMs, however, support many of the most popular cloud frameworks. In the case of some cloud offerings, like AWS with the Greengrass framework, an application running on the cloud can be seamlessly mobilized to the edge, or vice versa. Given the expectation of time and durability for most Industrial assets, Kria SOMs are made to adapt over the Industrial life cycle.

Getting Started with Kria SOMs

As mentioned throughout this white paper, Kria SOMs are the ideal solution for embedded and AI software developers that want to adopt the benefits of FPGA technologies in production systems but do not know where to begin. Kria SOMs also make Adaptive SoC hardware design more efficient and lower risk for FPGA experts by offloading some of the more time-consuming elements of the design process.

The Kria SOM developer ecosystem will continue to grow as Xilinx and partners endeavor to deliver additional accelerated applications and enhance the Kria family of SOMs and starter kits to exponentially decrease the time to harness the performance and flexibility of Xilinx Adaptive Computing. Designers can take the first step today by evaluating the current Kria SOM portfolio on the web at https://www.xilinx.com/products/som/kria.html. Documents, videos, and training material are available to assist with getting started. Designers can purchase a starter kit and an accessory pack, take a look at the Kria accelerated applications on the Xilinx App Store, and then identify the one best suited to theirs.

Acknowledgment

The following Xilinx employee has authored or contributed to this white paper:
Chetan Khona, Director, Industrial, Vision, Healthcare & Sciences.

Revision History

The following table shows the revision history for this document:

Date	Version	Description of Revisions
04/20/2021	1.09	Initial Xilinx release.

Disclaimer

The information disclosed to you hereunder (the "Materials") is provided solely for the selection and use of Xilinx products. To the maximum extent permitted by applicable law: (1) Materials are made available "AS IS" and with all faults, Xilinx hereby DISCLAIMS ALL WARRANTIES AND CONDITI/ONS, EXPRESS, IMPLIED, OR STATUTORY, INCLUDING BUT NOT LIMITED TO WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, OR FITNESS FOR ANY PARTICULAR PURPOSE; and (2) Xilinx shall not be liable (whether in contract or tort, including negligence, or under any other theory of liability) for any loss or damage of any kind or nature related to, arising under, or in connection with, the Materials (including your use of the Materials), including for any direct, indirect, special, incidental, or consequential loss or damage (including loss of data, profits, goodwill, or any type of loss or damage suffered as a result of any action brought by a third party) even if such damage or loss was reasonably foreseeable or Xilinx had been advised of the possibility of the same. Xilinx assumes no obligation to correct any errors contained in the Materials or to notify you of updates to the Materials or to product specifications. You may not reproduce, modify, distribute, or publicly display the Materials without prior written consent. Certain products are subject to the terms and conditions of Xilinx's limited warranty, please refer to Xilinx's Terms of Sale which can be viewed at http://www.xilinx.com/legal.htm#tos; IP cores may be subject to warranty and support terms contained in a license issued to you by Xilinx. Xilinx products are not designed or intended to be fail-safe or for use in any application requiring fail-safe performance; you assume sole risk and liability for use of Xilinx products in such critical applications, please refer to Xilinx's Terms of Sale which can be viewed at http://www.xilinx.com/legal.htm#tos.

Automotive Applications Disclaimer

XILINX PRODUCTS ARE NOT DESIGNED OR INTENDED TO BE FAIL-SAFE, OR FOR USE IN ANY APPLICATI/ON REQUIRING FAIL-SAFE PERFORMANCE, SUCH AS APPLICATI/ONS RELATED TO: (I) THE DEPLOYMENT OF AIRBAGS, (II) CONTROL OF A VEHICLE, UNLESS THERE IS A FAIL-SAFE OR REDUNDANCY FEATURE (WHICH DOES NOT INCLUDE USE OF SOFTWARE IN THE XILINX DEVICE TO IMPLEMENT THE REDUNDANCY) AND A WARNING SIGNAL UPON FAILURE TO THE OPERATOR, OR (III) USES THAT COULD LEAD TO DEATH OR PERSONAL INJURY. CUSTOMER ASSUMES THE SOLE RISK AND LIABILITY OF ANY USE OF XILINX PRODUCTS IN SUCH APPLICATI/ONS.