SDAccel™ Environment

The Xilinx SDAccel Development Environment

Bringing The Best Performance/Watt to the Data Center

Introduction

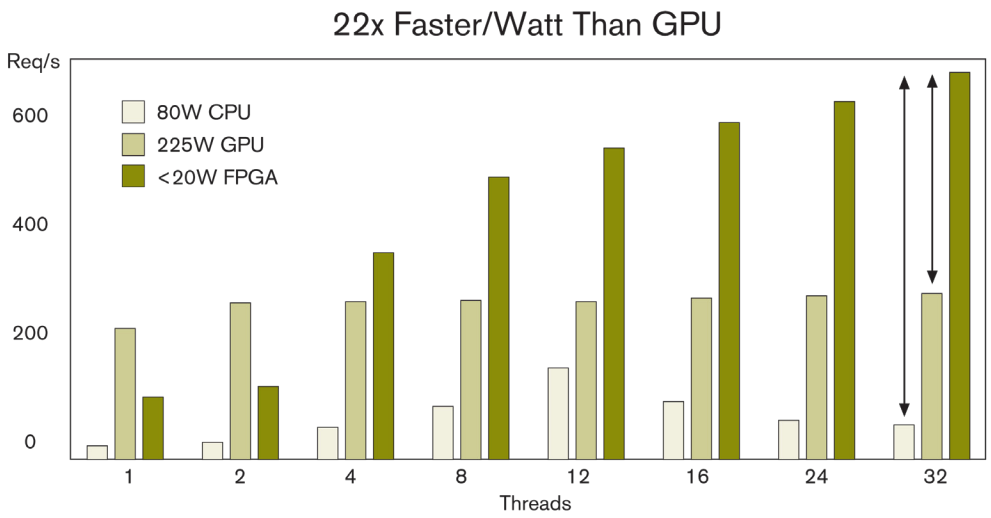
Data center operators constantly seek more server performance. Currently they develop applications with easy-to-program multicore CPUs and GPUs but CPU performance/watt is hitting the wall and GPU performance/watt is hitting the wall as well. Designers working on high-volume data-center applications including key acceleration, image recognition, speech transcription, encryption, and text search want GPU ease-of-programming but with hardware that will give them low power consumption, high throughput, and the lowest possible latency. However, there's a significant problem with scalability of multicore-CPU and GPU accelerators: developers would like to target simple full height plug-in PCIe® boards to use as application accelerators in data center servers. These boards can be configured to run high power graphics cards but customers want to target 25W or less to maximize scalability and minimize total power footprint.

Recent studies conducted jointly by Xilinx® and ETH Zurich, a Swiss university, have shown that FPGA-based application acceleration can achieve up to 25X better performance per watt and 50-75x latency improvement compared to CPU/GPU implementations while also providing excellent I/O integration (PCI, DDR4 SDRAM interfaces, high-speed Ethernet, etc.). In other words, FPGAs provide the heart of what's needed for power-efficient hardware application acceleration on one chip while providing solutions that are below the 25W per board targets.

Chinese web services leader Baidu® presented supporting findings at the 2014 Hot Chips Symposium held in San Jose, California. Baidu's conclusions:

- Mid-range FPGAs can achieve 375 GFLOPS dissipating only 10-20W
- FPGA-based accelerators can be deployed in all types of servers
- FPGAs can deliver more performance than CPUs or GPUs in a DNN (Deep Neural Network) prediction system. DNN systems are used for a variety of applications including speech recognition, image search, face recognition, web page search, and natural language processing.

XILINX ALL PROGRAMMABLE™



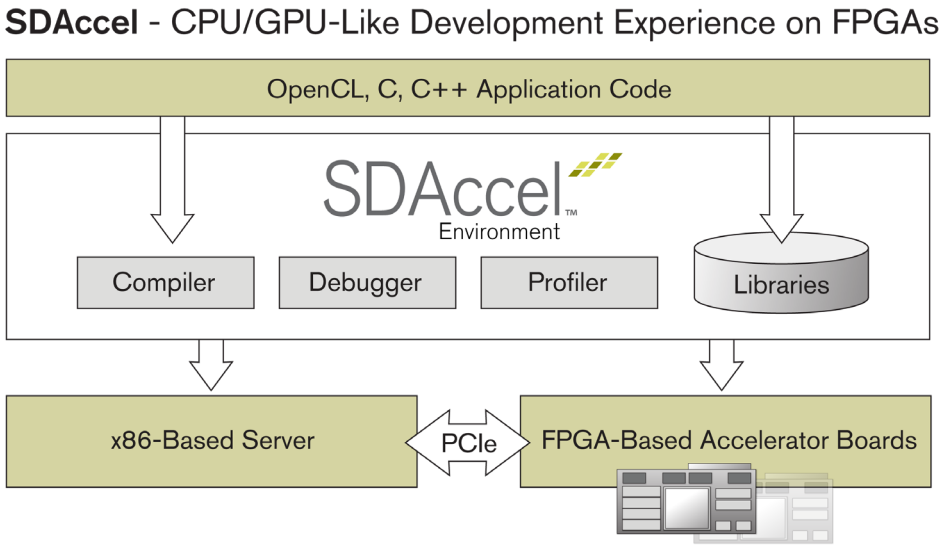
FPGAs provide significantly more hardware acceleration performance/watt

The biggest problem with FPGAs used for application acceleration has been programming. Data center application developers do not want to work with a hardware-centric RTL flow—the traditional FPGA development path. What they want is:

- High performance/watt
- A complete software development environment
- Easily upgradeable designs

Introducing the New Xilinx SDAccel Development Environment

The new Xilinx SDAccel™ Development Environment gives data center application developers the complete FPGA-based hardware and software solution they want. SDAccel includes a fast, architecturally optimizing compiler that makes efficient use of on-chip FPGA resources; a familiar software-development flow with an Eclipse™-based Integrated Design Environment (IDE) for code development, profiling, and debugging, which provides a CPU/GPU-like work environment; and dynamic reconfigurable accelerators optimized for different data center applications that can be swapped in and out on the fly for a CPU/GPU-like run-time environment. Applications can have many multiple kernels swapped in and out of the FPGA during runtime without disrupting the interface between the server CPU and the FPGA for nonstop application acceleration.



The SDAccel Development Environment includes an architecturally optimizing compiler, libraries, a debugger, and a profiler and provides a CPU/GPU-like programming experience

XILINX
ALL PROGRAMMABLE™

SDAccel combines the industry's first architecturally optimizing compiler, libraries, and development boards for the only CPU/GPU-like development and run-time experience using FPGAs. The SDAccel Development Environment targets host systems based on x86 server processors and provides commercial off-the-shelf (COTS) plug-in PCIe cards that add FPGA functionality.

First Architecturally Optimizing Compiler for C, C++, and OpenCL

SDAccel's architecturally optimizing compiler allows software developers to optimize and compile streaming, low-latency, and custom datapath applications. The SDAccel compiler targets high-performance Xilinx FPGAs and supports source code using any combination of OpenCL, C, C++, and kernels. The SDAccel compiler delivers as much as a 10X performance improvement over high-end CPUs and one tenth the power consumption of a GPU, while maintaining code compatibility and a traditional software programming model for easy application migration and cost savings. Based on partner benchmarks, the SDAccel compiler provides 3X the performance and resource efficiency of competitive FPGA solutions. The automatically generated designs from the SDAccel compiler can even deliver more performance than hand-coded RTL design solutions—as much as 20% in some cases.

The architecturally optimizing SDAccel compiler automatically applies a range of basic and advanced optimizations to the source application code. The following table lists the optimizations employed by the SDAccel compiler to accelerate applications.

Optimization	Description
Allocation	Tradeoff hardware resources and performance
Array Optimization	Array merging/partitioning/reshaping for performance, utilization
Data Packing (vector data type in ocl)	Pack data members to reduce ports and increase performance
Dataflow	Extract coarse-grain parallelism
Pipeline	Extract fine-grain parallelism
Balance	Balance expressions to improve performance
Loop Merge	Merge loops for performance and utilization
Loop Flatten	Flatten loop nests to improve performance
Loop Unroll	Unroll loops (partially or completely) to optimize performance
Stream (ocl: pipes)	Stream data to improve performance and utilization
Interface	Advanced interface synthesis
Resource	Resource selection for performance and utilization
Inline	Remove function boundaries to expose optimization opportunities
Instantiation	Function instantiation for performance
Dependence	Intra/inter dependence extraction
Bitwidthmin	Bit width minimization for performance and utilization
Reordered Math	Advanced floating point optimizations
ROM Inference	Turn complex computation s into ROMs when possible

Optimizations employed by the SDAccel compiler to accelerate applications.

Some of these automatic optimizations significantly improve overall application efficiency. For example, the SDAccel compiler's dataflow pipelining can pipeline multiple functions instead of being limited to optimizing single functions. The compiler's memory optimizations can greatly improve a function's efficiency so that you can fit more functions into the FPGA.

First CPU/GPU-Like Development Experience on FPGAs

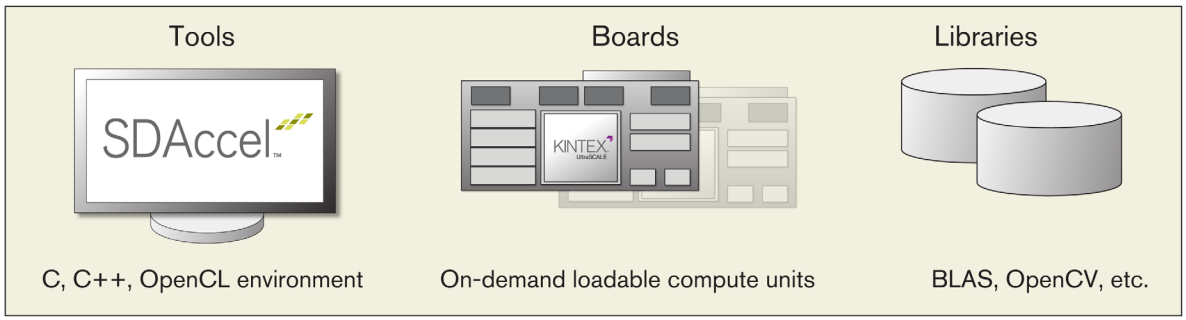
SDAccel is the first complete software development environment targeting FPGA platforms that enables a CPU/GPU-like development experience. Developers can use a familiar workflow to optimize their applications and take advantage of FPGA platforms with little to no prior FPGA experience. The IDE provides coding templates and SW libraries and enables compiling, debugging, profiling, and FPGA emulation on x86 platforms. When ready for deployment, it then implements the algorithm on data-center-ready COTS FPGA platforms complete with automatic instrumentation insertion.

XILINX
ALL PROGRAMMABLE™

SDAccel has also been architected to enable CPU/GPU developers to easily migrate their applications to FPGAs while maintaining and reusing their OpenCL, C, and C++ code in a familiar workflow. A large amount of code and libraries exist in C and C++ format. The ability to work in OpenCL, C, or C++ is a major benefit to developers who wish to leverage different code from their existing code base, use third party libraries, or even continue to develop in a mix of OpenCL, C, or C++ as the application warrants.

The SDAccel development environment is complete with the programmer-ready, Eclipse-based IDE, a host of C-based FPGA optimized libraries, as well as COTS platforms ready for data center use. Data-center-ready acceleration boards are available from Convey Computer, Alpha Data Parallel Systems, and Pico Computing. More COTS partners will be added early in 2015. SDAccel libraries include OpenCL built-ins, DSP, Video, and linear algebra libraries for high performance low power implementations. For domain-specific acceleration, Xilinx Alliance member Auviz Systems provides optimized OpenCV and BLAS SDAccel-compatible libraries.

SDAccel - A Programmer-Ready Platform



SDAccel: A Programmer-Ready Platform

Application developers can start to use SDAccel entirely in the x86 emulation space to get their code functional. When they are confident of their algorithms, they can profile the code to find code sections that would benefit from acceleration. Developers can then take these targeted sections and seamlessly use fast, automatically generated, cycle-accurate simulations of the accelerated kernels. These fast simulations can be used to debug and optimize the amount of acceleration while still working at an architectural level. The proven application is ready to port to the host/FPGA system. The SDAccel Development Environment supports all of these activities from a single cockpit.

SDAccel libraries contribute substantially to SDAccel's CPU/GPU-like development experience. These libraries include low-level math libraries and higher-productivity libraries such as BLAS, OpenCV, and DSP libraries. These libraries are written in C++ (as opposed to RTL), so that they can be used exactly as written during all development and debugging phases. Early in a project, all development will be done on the CPU host. Because the SDAccel libraries are written in C++, they can simply be compiled along with the application code for a CPU target—creating a virtual prototype—which permits all testing, debugging, and initial profiling to occur initially on the host. During this phase, no FPGA is needed.

Once the application has been initially debugged and profiled, the critical functions in need of hardware acceleration can be compiled for co-simulation, with the accelerated functions running on a CPU-based RTL simulator. At this point, the software development team can observe the resulting performance and verify that the application will meet performance objectives. RTL simulation gives a very accurate picture of the performance to be expected from FPGA hardware acceleration so this ability gives the development team a very good performance estimate and a high level of assurance that the produced FPGA configuration exactly reproduces the behavior and results of the application source code.

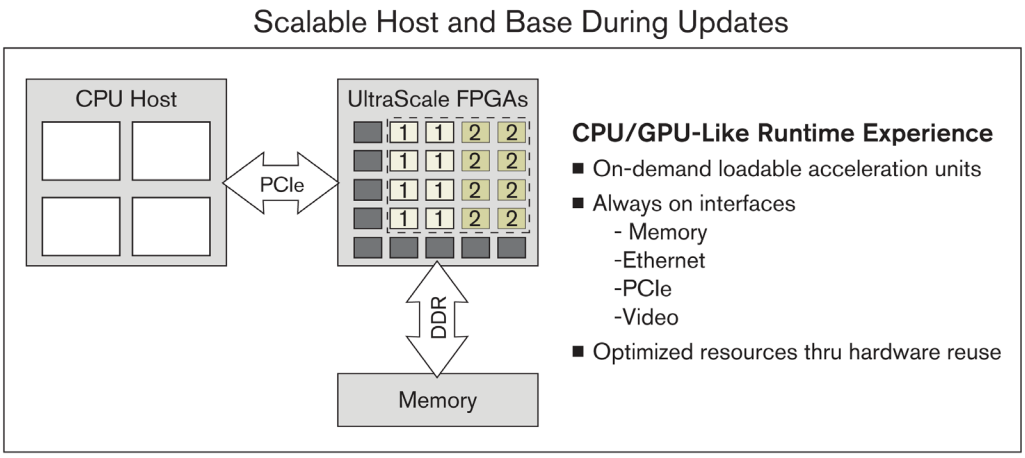
XILINX
ALL PROGRAMMABLE™

The accelerated functions can then be moved to the FPGA to produce the final accelerated application, ready for deployment. Instead of creating a generic FPGA configuration, the SDAccel compiler targets specific Xilinx FPGAs at compile time—including the device speed grade—so the resulting accelerated functions are optimized with respect to the available resources on the FPGA and the device timing. Even at this point, the functions can be instrumented—which imposes no performance penalty and very little hardware overhead—and additional late-stage profiling can occur within the same familiar SDAccel Development Environment to take fullest advantage of the FPGA hardware and to fully maximize application acceleration.

CPU-level profiling identifies the application bottlenecks. Co-simulation will tell you if the bottlenecks have been eliminated by accelerating specific functions. Profiling hardware accelerated functions running on the FPGA permits additional optimizations and further fine tuning of the application code and the accelerated functions. At all times and in all phases, the SDAccel compiler can optionally instrument the application code including the accelerated functions, which means that debugging and profiling can proceed in the same manner and using the same software-development tools no matter which phase the development is in.

First CPU/GPU-Like Run-time and Update Experience on FPGAs

SDAccel offers the only FPGA-based dynamic reconfigurable accelerators that enable real-time CPU/GPU-like run-time updates. Unique to FPGA solutions, SDAccel keeps the system functional during kernel updates with the only FPGA-based dynamic reconfigurable capability that can load new hardware accelerator kernels—similar to the abilities of CPU/GPU accelerators—while keeping critical system interfaces and functions such as memory, Ethernet, PCIe, and performance monitors live. This on-the-fly system reconfiguration is ideal for immediate updates to data center compute needs and loads. An example of an application where this ability is of strategic advantage: switching between image search, video transcoding, and image processing on the fly.



Reconfigurable accelerators permit rapid hardware acceleration changes, maximizing accelerator ROI

All of this translates into resource optimization through hardware reuse, which is a significant advantage in data center environments. In simple terms, an SDAccel-based system can accelerate one application today and if another type of acceleration is needed tomorrow, the system can be upgraded quickly and smoothly.

Developers can manage and run SDAccel-based systems just as they would a CPU/GPU system. With SDAccel and this rapid dynamic reconfiguration of the hardware, it is even possible to reuse the FPGA while the application is running. This capability allows developers to perform multiple different kernel accelerations using the same FPGA board with the hardware being re-optimized for each algorithm without restricting or interrupting the application flow. By permitting this dynamic reconfiguration, SDAccel allows data center operators to get the most out of their hardware-acceleration investment.

XILINX
ALL PROGRAMMABLE™

Concrete Benchmarks Prove SDAccel's Performance and Value

SDAccel is industry proven with real world benchmarks. For example, in a compression benchmark, SDAccel produced hardware accelerators that are over 3x faster and over 3x smaller than accelerators generated by a competing tool. Area and throughput is comparable or better than hand optimized RTL implementations.

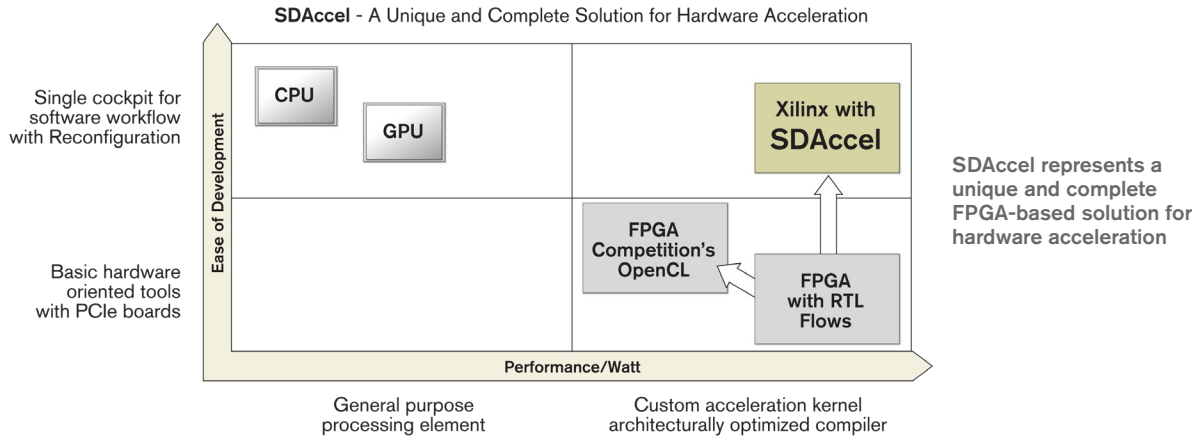
	Metrics	Hand Coded RTL	SDAccel Environment	Other FPGA OpenCL SW
Compression Benchmark	Throughput	1x	1x	0.3x
	Area	1x	0.9x	3x
Encryption Benchmark	Throughput	1x	1.2x	1x
	Area	1x	1.25x	5x

Two accelerator benchmarks comparing hand coded RTL, SDAccel and results generated by a competing tool

For the above encryption benchmark, SDAccel produced hardware accelerators that are 1.2x faster and 4x smaller than accelerators generated by a competing tool. Even better, the accelerator automatically generated by SDAccel was 20% faster than an accelerator that had been hand-coded using RTL. Out of the box, SDAccel can compile accelerators with hand-coded QOR (quality of results) or better.

Conclusion

SDAccel gives data center application developers what they want: a CPU/GPU-like work environment that produces efficient hardware accelerators with the best available performance/watt—far superior to accelerators running on CPU/GPU servers. SDAccel supports a software workflow through a single cockpit environment with in-system, on-the-fly reconfigurability that maximizes hardware acceleration ROI in the data center. By doing so, SDAccel represents a unique and complete FPGA-based solution—a solution that far exceeds the ease of use and the abilities of competing point-tools, which cannot offer the same high-productivity level of workflow and environment familiarity and cannot match the superior performance of hardware accelerators automatically generated by SDAccel.



Take the NEXT STEP

Please visit: **www.xilinx.com/sdaccel**

To take advantage of the unique performance/watt advantages of FPGA-based hardware acceleration, please contact a local Xilinx sales representative for more information about Xilinx SDAccel.

Corporate Headquarters
Xilinx, Inc.
2100 Logic Drive San Jose, CA
95124 USA
Tel: 408-559-7778
www.xilinx.com

Europe
Xilinx Europe One Logic Drive
Citywest Business Campus Saggart,
County Dublin Ireland
Tel: +353-1-464-0311
www.xilinx.com

Japan
Xilinx K.K.
Art Village Osaki Central Tower 4F
1-2-2 Osaki, Shinagawa-ku
Tokyo 141-0032 Japan
Tel: +81- 3-6744-7777
japan.xilinx.com

Asia Pacific Pte. Ltd.
Xilinx, Asia Pacific
5 Changi Business Park
Singapore 486040
Tel: +65-6407-3000
www.xilinx.com

India
Meenakshi Tech Park
Block A, B, C, 8th & 13th floors,
Meenakshi Tech Park, Survey No. 39
Gachibowli(V), Seri Lingampally (M),
Hyderabad -500 084
Tel: +91-40-6721-4000
www.xilinx.com

XILINX
ALL PROGRAMMABLE™