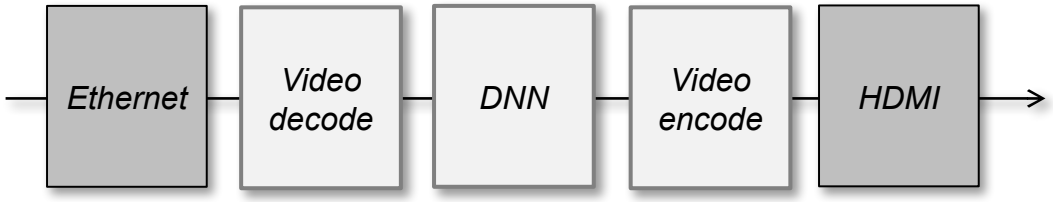
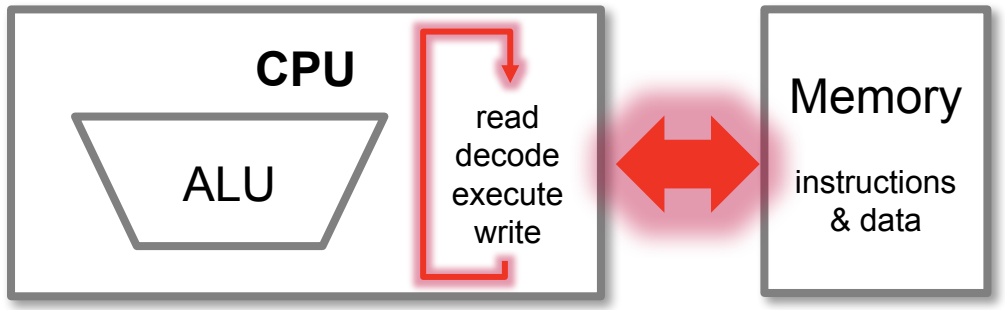
Xilinx and Amazon EC2 F1 Compute Instance

March 23, 2018

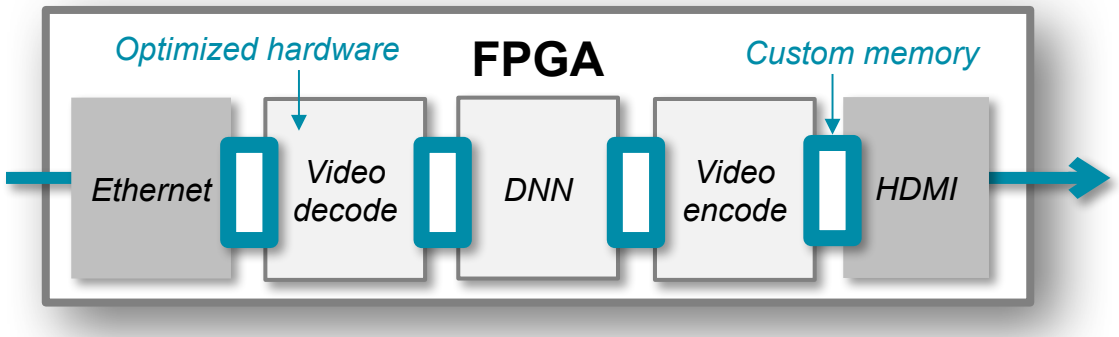
Why FPGA: Application Specific HW & Memory



- Given an algorithm to implement…



- CPU/GPU implementation: Von Neumann
  - Rigid sequential execution (SIMD for GPU)
  - Memory access bottleneck
  - Not optimal for custom width or decision handling



- FPGA implementation: WYSIWYG
  - **Custom dataflow**, width, decision handling
  - **Custom memory hierarchy**: keeps data inside
  - **Custom IOs**: high throughput & low latency

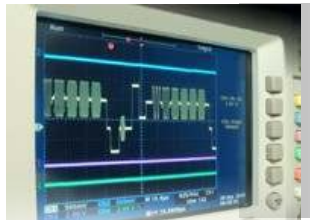
Where are FPGAs Traditionally Used?

Communications
- Wired networking
- Wireless infrastructure

Automotive
- Infotainment
- Driver assistance

Datacenter
- High performance computing
- Solid state drives

Aerospace and Defense
- Avionics, Communications
- Space



Test & Measurement
- Communications instruments
- Semiconductor test equipment

Audio, Video, Broadcast
- 3D cameras
- Video transport

Industrial, Scientific, Medical
- Ultrasound systems
- Motor controllers

Consumer
- 3D television
- eReaders

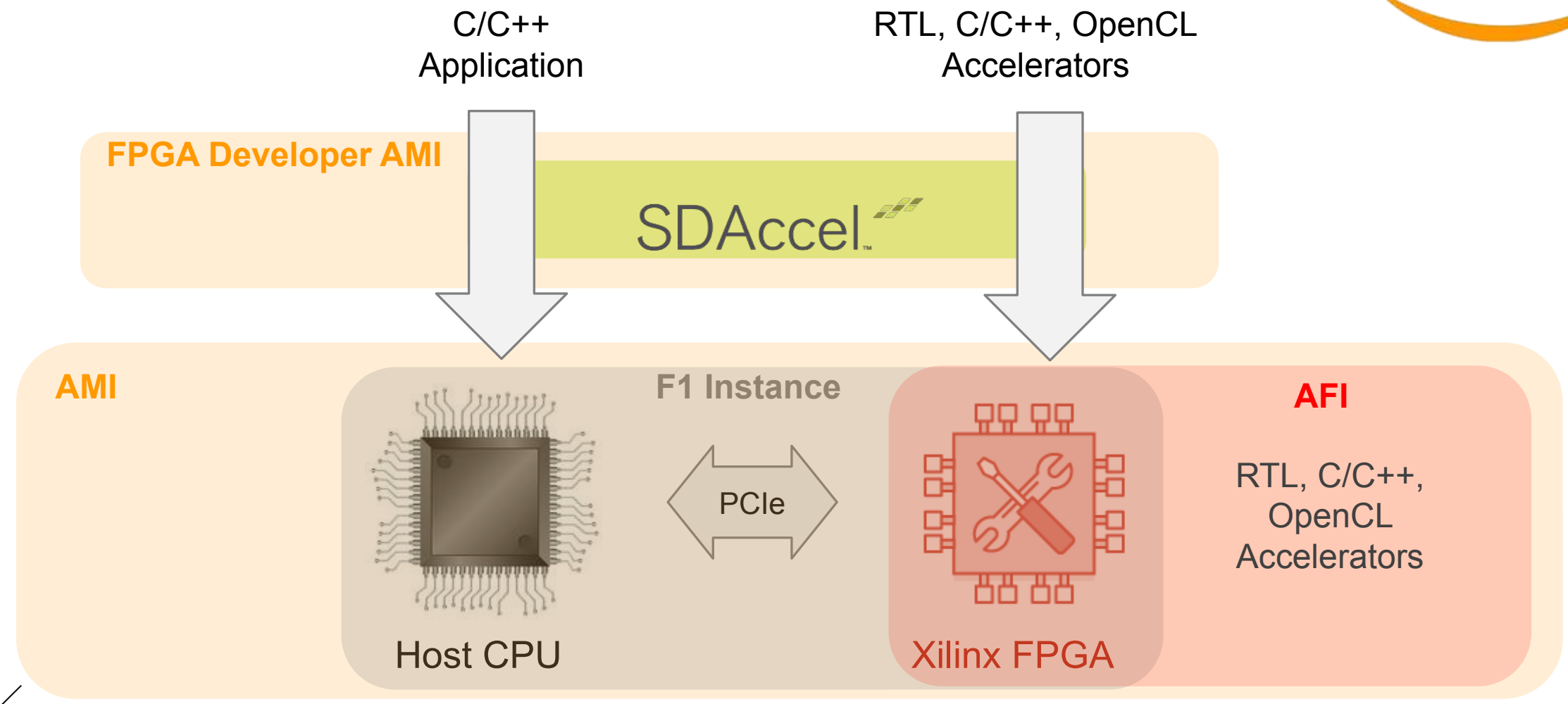
Introducing the AWS EC2 F1 Instance

EC2 F1 instance is a compute instance based on Xilinx FPGA. It can be reconfigured as an application-specific HW accelerator

FPGA Developer AMI makes the Xilinx SDAccel environment available to simulate, compile, build and register Amazon FPGA Images (AFI)

Amazon FPGA Image (AFI) is the design to program the FPGA. AWS provides a service to manage AFIs. AFIs can be associated with an AMI and offered in the AWS Marketplace

Introducing the AWS EC2 F1 Instance



awsmarketplace

XILINX ALL PROGRAMMABLE.

Benefits of AWS F1

- Business model
  - Turns expensive HW appliance to cost effective, pay-per-use SaaS / API
  - Takes sales cycles / evaluations from months to hours
  - Scales to millions of AWS users

- Development model
  - Access to Cloud-based SW development tools
  - Access to fastest / best FPGA from anywhere
  - Applications can elastically scale to as many accelerators as needed

Amazon EC2 F1 Applications

aws marketplace

AMI & SaaS

FPGA Developer AMI
(2) | Version 1.3.3 | Sold by Amazon Web Services
The FPGA (field programmable gate array) AMI is a supported and maintained CentOS Linux image provided by Amazon Web Services. The AMI is pre-built with FPGA development...
Linux/Unix, CentOS 7.3 - 64-bit Amazon Machine Image (AMI)

ZEBRA on 1 FPGA (image classification)
(0) | Version 2017.04.1 | Sold by Mipsology
Zebra offers users FPGA-based class-leading acceleration for neural network inference. The user-defined neural network works on Zebra just as it would on GPU or CPU. Zebra...
Linux/Unix, CentOS 7.3 - 64-bit Amazon Machine Image (AMI)

Zebra Deep-Learning engine for Caffe (1 FPGA)
Free Trial
(0) | Version V2017.10.1 | Sold by Mipsology SAS
Starting from **$0.10** to **$0.10/hr** for software + AWS usage fees
Zebra accelerates neural network inference using FPGA. User-defined neural networks are computed by Zebra just as they would be by a GPU or a CPU. Zebra is fully integrated...
Linux/Unix, Ubuntu 14.04 - 64-bit Amazon Machine Image (AMI)

FPGA-Accelerated Deep-Learning Inference with Binarized Neural Networks
(0) | Version 1.2 | Sold by Missing Link Electronics, Inc.
Starting from **$5.00** to **$5.00/hr** for software + AWS usage fees
Image classification of the Cifar10 dataset using the CNV neural network. Based on Xilinx public proof-of-concept implementation of a reduced-precision, Binarized Neural Network...
Linux/Unix, Ubuntu 16.04 - 64-bit Amazon Machine Image (AMI)

Visual System Integrator for FPGA and Embedded Development
(0) | Version 2017.1_Autoupdate | Sold by System View
Starting from **$0.50/hr** or from **$2,500.00/yr** (43% savings) for software + AWS usage fees
Visual System Integrator is the one-of-a-kind tool for embedded development which, for the first time, makes it possible to develop a full functioning system. Visual System...
Linux/Unix, CentOS 7.3 - 64-bit Amazon Machine Image (AMI)

Hyperion F1 10G RegEx File Scan
(0) | Version 5.5.2 | Sold by TITAN IC SYSTEMS LTD
Starting from **$50.00** to **$50.00/hr** for software + AWS usage fees
The Hyperion F1 10G RegEx File Scan instance provides a preloaded IP image of the 10Gbs Regular eXpression Processor (RXP) on the FPGA and a pre-installed SDK allowing the...
Linux/Unix, Amazon Linux 2017.09.0 - 64-bit Amazon Machine Image (AMI)

FireSim Demo v1.0
(0) | Version 1.0 | Sold by Berkeley Architecture Research
FireSim is an FPGA-accelerated hardware simulation tool that cycle-accurately simulates RISC-V RocketChip-based clusters, with peripherals like disks and network interface...
Linux/Unix, CentOS 7.3 - 64-bit Amazon Machine Image (AMI)

Merlin Compiler AMI
(0) | Version 1.0.1a | Sold by Falcon Computing Solutions, Inc.
14 Day Free Trial Available - The Merlin Compiler AMI is provided by Falcon Computing Solutions, Inc. The AMI is pre-built with Merlin Compiler that provides push-button C/C++...
Linux/Unix, CentOS CentOS 7.3 - 64-bit Amazon Machine Image (AMI)

InTime
(0) | Version 2.3.0 | Sold by Plunify
InTime is an automated optimization software for FPGA design by Plunify. It optimizes timing and design performance using machine learning to find the best combination of...
Linux/Unix, CentOS 6.9 - 64-bit Amazon Machine Image (AMI)

NGCodec HEVC/H.265 Encoder D01
Free Trial
(0) | Version 5.01 | Sold by NGCodec
Starting from **$1.25** to **$10.00/hr** for software + AWS usage fees
Using an F1 instance, offload HEVC encoding to an FPGA. This version of the NGCodec Encoder features I&P&B frame encoding at up to 1080p60 resolution/frame rate. The performance...
Linux/Unix, CentOS 7 - 64-bit Amazon Machine Image (AMI)

Machine Learning Development Stack from Xilinx, Preview Edition
(0) | Version 17.12.15 | Sold by Xilinx
In this Machine Learning Development Stack, Preview Edition AMI,users easily integrate machine learning into their current applications and deploy them quickly. Users can...
Linux/Unix, CentOS 7.3 - 64-bit Amazon Machine Image (AMI)

Toolkit Powered by RYFT Heterogeneous Computing
(0) | Version v2.1.1 | Sold by Ryft
Starting from **$5.00** to **$5.00/hr** for software + AWS usage fees
Ryft's Toolkit is a pre-configured, ready to run image for instantly integrating smarter, more sophisticated FPGA-accelerated search & analysis capabilities into existing...
Linux/Unix, Ubuntu 16.04 - 64-bit Amazon Machine Image (AMI)

Elasticsearch Powered By RYFT Heterogeneous Computing
(0) | Version v2.1.1 | Sold by Ryft
Starting from **$8.00** to **$8.00/hr** for software + AWS usage fees
Ryft's ELK is a pre-configured, ready to run image for deploying the powerful open source, distributed real-time search and analytics engine, Elasticsearch, on Amazon's FPGA-accelerated...
Linux/Unix, Ubuntu 16.04 - 64-bit Amazon Machine Image (AMI)

Torch: A scientific computing framework for LuaJIT by Miri Infotech
Free Trial
(0) | Version 1* | Sold by Miri Infotech
Starting from **$0.03** to **$0.03/hr** for software + AWS usage fees
Torch is a suite of business tools that uses data mining, machine learning and artificial intelligence to automate work, save money and helps the business grow.It is easy...
Linux/Unix, Amazon Linux 2016.03.03 - 64-bit Amazon Machine Image (AMI)

DeePhi Descartes Efficient Speech Recognition Engine
(0) | Version 2018.02.2a | Sold by Beijing DeePhi Technology Co., Ltd.
This is an end-to-end ASR (Automatic Speech Recognition) system with FPGA acceleration on AWS F1 by DeePhi. We modify the Baidu DeepSpeech2 framework (https://github.com/SeanNaren/deepspeech.pytorch)...
Linux/Unix, CentOS 3.10.0-693.2.2.el7.x86_64 - 64-bit Amazon Machine Image (AMI)

Accelerated Machine Learning
Free Trial
(0) | Version AML_v1.0 | Sold by InAccel
Starting from **$2.00** to **$2.00/hr** for software + AWS usage fees
AML is InAccel's accelerated machine learning library. It aims to maintain the practical and easy to use interface of other open source frameworks, i.e. of Apache Spark, and...
Linux/Unix, Ubuntu 16.04 - 64-bit Amazon Machine Image (AMI)

DRAGEN Complete Suite - Exome (approx. $2 per Exome)
Free Trial
(0) | Version 2.2 | Sold by Edico Genome
Starting from **$9.25** to **$22.80/hr** for software + AWS usage fees
The DRAGEN Complete Suite (Exome) enables ultra-rapid analysis of Next Generation Sequencing (NGS) data for small data sets, such as whole exomes and targeted panels. This...
Linux/Unix, CentOS 7.2 - 64-bit Amazon Machine Image (AMI)

DRAGEN Complete Suite - Genome (approx. $15 per Genome)
Free Trial
(0) | Version 2.2 | Sold by Edico Genome
Starting from **$10.35** to **$22.10/hr** for software + AWS usage fees
The DRAGEN Complete Suite (Genome)* enables ultra-rapid analysis of Next Generation Sequencing (NGS) data for large data sets, such as whole genomes. This application uses...
Linux/Unix, CentOS 7.2 - 64-bit Amazon Machine Image (AMI)

Compelling Use Applications on Amazon EC2 F1

Machine Learning Inference
Speech recognition

40x

DEEPHI 深鉴科技

Video Streaming
Frame rate for HEVC encoding

10x

NGCODEC NEXT GENERATION VIDEO COMPRESSION

Genomics
20 min vs. 33 hours for whole genome analysis

100x

edico genome

Big Data Analytics
40 min vs. 60 hours for logfile query

90x

RYFT™ ACTIONABLE INTELLIGENCE FROM COMPLEX DATA

Example: Enabling Personalized Medicine

edico genome

- Genome diagnosis to treat critically ill newborns
- Analytics reduced from 1+ day to 20 minutes
- Dynamically reconfigures for patient-specific genomics acceleration

DRAGEN Genome Exome | DRAGEN Methylation Epigenome | DRAGEN Microbiome Metagenome | DRAGEN Joint Genotyping | DRAGEN Transcriptome RNAseq | DRAGEN Cancer Toolkit | DRAGEN Advanced Variant Caller | DRAGEN Structural Variant /CNVs | gatk GATK Best Practices

Xilinx Cloud Partner Program

Accelerator Program

A quick start program to enable companies to accelerate products and services with FPGAs on the cloud. Target workloads include data analytics, genomics, video processing, machine learning, financial technology, security, and storage.

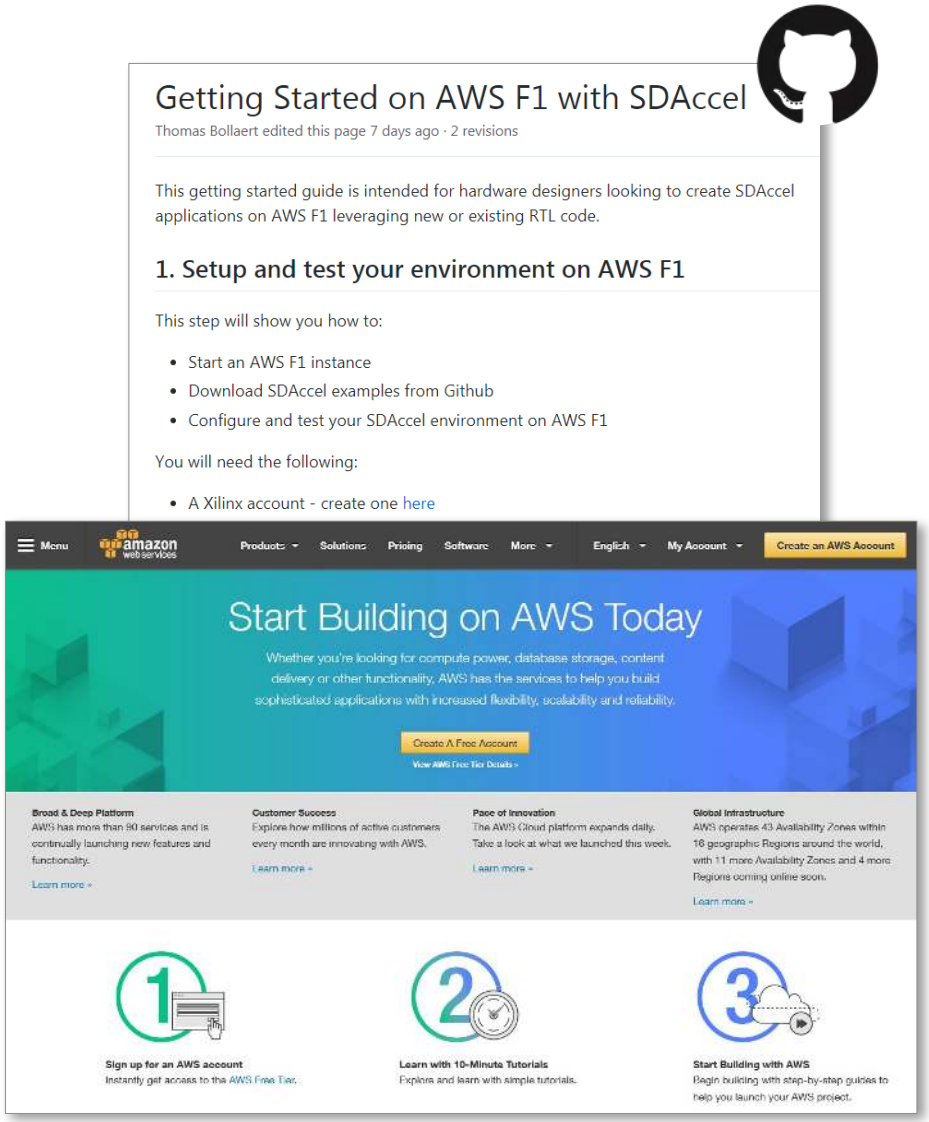
Apply Now

https://www.xilinx.com/products/design-tools/acceleration-zone/accelerator-program.html

Get Started on AWS F1 in 5 Simple Steps

- Onboarding program with step-by-step instructions, online training and Github examples

1. Setup and test your AWS environment
2. Run your first SDAccel hello world!
3. Build your SDAccel knowledge
4. Practice and Experiment
5. Install and run SDAccel on your own machine



https://github.com/Xilinx/SDAccel_Examples/wiki/Getting-Started-on-AWS-F1-with-SDAccel-and-RTL-Kernels