State of FPGA-based Acceleration

Presented By

Vinay Singh
Sr. Director, Software Solutions and Platform Marketing

October 2, 2018
FPGA-Based Acceleration Momentum

Accelerated Applications:
- 250+ Companies Developing Applications

Developers:
- 10x More Developers Onboarded since 2017

Acceleration Platforms:
- 4 Cloud
- 2 On-premise
Develop & Deploy Applications From Cloud to On-Premise

F1 Instance VU9P  |  F1 Instance VU9P  |  FP1 Instance VU9P  |  F3 Instance VU9P  |  New instance F1.4xlarge

AWS GovCloud (US)  |  HUAWEI  |  Aliyun  |  AWS  |  Nimbix
AWS  |  Alibaba Cloud Computing  |  AWS  |  Supercomputing made super simple

On-premise
Alveo U200, U250  |  Alveo U200, U250
Oct. 2, 2018  |  Oct. 2, 2018

General Availability
© Copyright 2018 Xilinx
Announcing EC2 F1 New Capabilities

• New instance size: f1.4xlarge provides a significant performance boost compared to f1.2xlarge

• Virtual Ethernet: enabling high-performance networking acceleration use-cases like firewalls, routers, filtering and more

• DRAM data retention: boosting FPGA images pipeline runtime execution

• New FPGA developer AMI supporting Vivado 2018.2 for faster compile times, higher frequencies and improved timing closure

> → Visit AWS Developer Hangout zone to learn more!
## Acceleration Ecosystem

<table>
<thead>
<tr>
<th>Developers</th>
<th>Partners</th>
<th>Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>115</strong></td>
<td><strong>43</strong></td>
<td><strong>15</strong></td>
</tr>
<tr>
<td>2017</td>
<td>2017</td>
<td>2017</td>
</tr>
<tr>
<td>Developers Trained Companies + Academia</td>
<td>Accelerator Program (Approved Partner)</td>
<td>Published Apps</td>
</tr>
<tr>
<td><strong>1127</strong></td>
<td><strong>268</strong></td>
<td><strong>35</strong></td>
</tr>
<tr>
<td>2018</td>
<td>2018</td>
<td>2018</td>
</tr>
</tbody>
</table>

© Copyright 2018 Xilinx
Getting Developers Onboarded

In-person

Co-located at strategic events
SC17, FPGA2018, XDF, CVPR, …

Virtual on-demand

Amazon EC2 F1 SDAccel Developer Lab

Work through this self-paced tutorial where you will receive an overview of AWS F1 and SDAccel™ with step-by-step instructions on using Amazon EC2 F1 instances to accelerate your applications. In this virtual developer lab, you will connect to an F1 instance, experience F1 acceleration, and develop and optimize F1 applications with SDAccel.

https://www.xilinx.com/products/design-tools/cloud-based-acceleration.html#sdaccellab

Developers Trained
2017 115
2018 1127

Companies + Academia

University Courses

Prof. Zhiru Zhang, Cornell
High-level Digital Design Automation Labs and project offered on AWS F1, Fall 2018

Prof. Jason Cong, UCLA
Customizable Computing for Big Applications, Fall 2018
Parallel and Distributed Computing, Winter 2019

Prof. Krste Asanovic, UC Berkeley
CS152 Computer Architecture and Engineering, Spring 2018

eLearning & Instructor-Led Courses

© Copyright 2018 Xilinx
Acceleration In Education

Xilinx University Program

Acceleration Your Research on the AWS Cloud with Xilinx FPGAs

- Scalable infrastructure
- No tools or license setup
- Integrate with over 100 AWS services
- Get started with AWS ECC F1 instances

www.xilinx.com/xup

Open Hardware design competition EMEA 2018

AWS F1 Category Winner and Finalists

5 Points to Rule Them All
Winner: AWS EC2 F1 prize

Implementation
Optimize the algorithm for the FPGA, preserving precision accuracy and adaptability of the software version

The Solution
Politecnico di Milano
Emanuele Del Sasso, Marco Rebozzi, Lorenzo Di Tucci
(Prof. Marco D. Santambrogio)
Project Link

AWS EC2 F1: circFA
Politecnico di Milano
Emanuele Del Sasso, Marco Rebozzi, Lorenzo Di Tucci
(Prof. Marco D. Santambrogio)
Project Link

AWS EC2 F1: FastBrain
Politecnico di Milano
Filippo Caroli, Giada Casagrande, Valentina Corbetta
(Prof. Marco Santambrogio)

Acceleration Posters @ XDF

FireSim: Productive, Scalable, FPGA-Accelerated Cycle-Accurate Hardware Simulation using Cloud FPGAs
Susan Kaeding, Howard Mills, Vincent Rice, Chris Burckel, Jess Neal, Ben Fullard, David Hare, Michael J. Reed, Christian Bader, John Andrade, Young Jee Piao, Christopher Chiellini, and Brian Van de Velde

Automata Processing on FPGAs
Chuan Guo, Vinh Dong, Ted Xin, Jack Wadden, Murat Shan, Kevin Shadron
Department of Computer Science, University of Virginia, structure@virginia.edu.

DEMOCRATIZE ACCELERATED GENOMIC PIPELINES ON FPGA
Lorenzo Di Tucci, Giulia Guidi, Marco Rebozzi, Sara Notargiacomo, Marco D. Santambrogio
Dipartimento di Informatica, Informazione e Scienze della Computeristica, Politecnico di Milano (Italy), santamb@polimi.it

GalaPagos: A Full Stack Approach to FPGA Integration in the Cloud
Naif Tarafdar, Nariman Eskandari, Paul Chow
Electrical and Computer Engineering, University of Toronto, naif.tarafdar@gmail.com

FPGAACCELERATED BASEBAND FOR WIRELESS SYSTEMS
Chaoxing Tang, Kaiyang Li, Joseph Cavallaro
Department of Electrical and Computer Engineering, New Mexico State University, canche@nmsu.edu

Rosetta: A Realistic HLS Benchmark Suite for FPGAs
Students: Yuan Zhou, Ritchie Zhuo, Hunchen Jin
Faculty Advisor: Zhuo Zhang
School of Electrical and Computer Engineering, Georgia Institute of Technology

www.inaccel.com
xeler.io/

Accelerator Startups from Academia

© Copyright 2018 Xilinx
Xilinx Accelerator Program for Developers

Enable companies to accelerate products/services on the cloud and on-premise

- Technical Enablement
- GTM Support
- Driving Revenue

- Leverage your Existing algorithms and IP
- Deploy Custom Hardware to Millions in Public Cloud
- Extend your existing business model
- Discounted XBB cards for development
- Connect to Xilinx marketing, FPGA expertise and funding opportunities

www.xilinx.com/accelerator-program
Accelerated Applications

- Compression
  - CAST
  - Accelize
  - XILINX

- Data Analytics
  - BLACKLYNX
  - XILINX
  - LegUp
  - BigZetta Systems
  - Axonerve

- Published Apps
  - 2017: 15
  - 2018: 35

- Financial Computing
  - MAXELER Technologies
  - SUMUP ANALYTICS
  - Politecnico di Torino

- Genomics
  - deLabiab Genetics
  - edico genome

- Image Processing
  - XILINX
  - CTACCEL
  - Accelerated Computing

- Machine Learning
  - DEEPHI
  - XILINX
  - inacel
  - mle
  - Mipsology

- Security
  - ZO TECH
  - Titan
  - SECURE-iC

- Video
  - NGCODEC
  - PATHPARTNER
  - skreens

- Tool
  - XILINX PLUNIFY
  - FireSim

https://www.xilinx.com/products/design-tools/acceleration-zone.html#libraries
Xilinx ML Suite - Fastest Real Time Inference

Your Application

{RESTful API}  python™

TensorFlow  mxnet  Caffe

xfdNN  Compiler • Auto-quantizer ML Software Libraries

XILINX

GPU

CPU

0  1000  2000  3000  4000

Googlenet v1  Img/s

* See White Paper for performance details

https://www.xilinx.com/ml
https://github.com/xilinx/ml-suite

© Copyright 2018 Xilinx
Xilinx ABR Video - Real Time Video Streaming

High performance HEVC and VP9 Encoder
- Fully configured transcoding pipeline
- Easy programming interface with FFmpeg

60 fps for real-time video streaming
- 7x greater than x265 slow
- 10x greater than Libvpx

© Copyright 2018 Xilinx
## Acceleration Resources

<table>
<thead>
<tr>
<th>Developers</th>
<th>Data Scientists, ML Practitioners</th>
<th>Academics</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWS F1 SDAccel Developer Lab</td>
<td>AWS F1 ML Suite Developer Lab</td>
<td>Sign-up for: AWS Educate, Xilinx University Program</td>
</tr>
<tr>
<td>Start-ups, IP Providers</td>
<td>Cloud End Users</td>
<td>On-premise End Users</td>
</tr>
<tr>
<td>Sign-up for: Xilinx Accelerator Program</td>
<td>AWS F1 Apps and Libraries</td>
<td>Xilinx Alveo Accelerator Cards</td>
</tr>
</tbody>
</table>
FPGA-based Acceleration Momentum

Accelerated Apps

Developers

Acceleration Platforms

© Copyright 2018 Xilinx