Xilinx Vision & Strategy for the Adaptable World

> Victor Peng, President & CEO
Three Big Trends

01

Explosion of Data

> 90% unstructured
> Video & image content
> Needs higher throughput & real-time computing
Three Big Trends

02

Dawn of AI

> Adoption across all industries
> Injecting new intelligence into apps
> From endpoints to edge to cloud
Three Big Trends

Computing After Moore’s Law

> Heterogeneous computing with accelerators
> Breadth of apps require different architectures
> Speed of innovation outpacing silicon design cycles
The intelligent connected world needs adaptable accelerated computing.
Strategy for Enabling the Adaptable World
Strategy for Enabling the Adaptable World

Data Center First
Reaching Software Application Developers

Accelerated Open Frameworks
- TensorFlow™
- gatk
- FFmpeg

Accelerated Libraries
- Machine learning
- H.265 HEVC
- Database analytics
- OpenCV

Development Environment
- SDAccel Environment
- SDSoc Environment

Development Boards

Software Application Developers

Development Stack

System Developers
Growing Data Center Compute Ecosystem

Applications, Tools & Communities
bitfusion, DEEPHi, edico genome, NGCODEC, RYFT

Cloud Development & Deployment (FPGA as a Service)
aws, Alibaba Cloud, Huawei, Baidu, Nimbix, Tencent Cloud

Technology & Systems
CCIX, AMD, arm, IBM, Mellanox Technologies, Qualcomm
Compute Acceleration

- Machine Learning 40x*
- Video & Image Processing 10x*
- Data Analytics 90x*
- Genomics 100x*
Genomics Use Case: Personalized Medicine

Whole genome diagnosis to treat critically ill newborns

Analysis reduced from 1 day to 20 minutes

Patient-specific genomics dynamically optimized
Computational Storage & Network Acceleration

Computational Storage

SmartNICs & Network Acceleration
Accelerate Growth in Core Markets
Accelerate Growth in Core Markets

- Automotive
- Wireless Infrastructure
- Wired Communications
- Audio, Video, & Broadcast
- Aerospace & Defense
- Industrial, Scientific & Medical
- Test, Measure, & Emulation
- Consumer
Drive Adaptive Computing
From FPGA to Adaptive Compute Acceleration Platform

ACAP

New Device Category for Adaptive Workload-Specific Acceleration

> HW/SW programmable engines

> IP subsystems and a network-on-chip

> Highly integrated programmable I/O
Benefits of an Adaptive Compute Acceleration Platform

Dynamically Adaptable to Workloads

Exponential Increase in Acceleration

Software Programmable

Custom acceleration for any workload – in milliseconds
Project “Everest”

The First 7nm ACAP Product Family
Project “Everest”

4 Years 1,500 Engineers 50B Transistors >$1B R&D
Everest Breakthroughs vs Current Generation

**Revolutionary Adaptability**
Dynamic Optimization for Workloads

**Software & Hardware Users**
Rapid Innovation & Deployment

**20x**
AI Compute Capability

**4x**
5G Communications Bandwidth
Project “Everest”

Timeline

> Early software tools available to strategic customers
> Silicon tape out this year
> First shipment next year
Vision - Adaptable Intelligent World

Strategy
Data center first
Accelerate growth of core vertical markets
Drive adaptive computing

ACAP – a new product category
Accelerating broad range of workloads with dynamically adaptable silicon
10-100x faster than CPUs for new workloads, more use cases than GPUs or ASICs

Project “Everest”
1st ACAP implemented in 7nm, tape out 2018
SW and HW programmable
>10X in performance, performance/watt
Thank you
Footnotes


**20X AI Compute** is based on a NIC / Data Center comparison with Virtex UltraScale+ for Machine Learning inference for image recognition. Everest compute is equivalent to 20x VU9Ps (Data Center’s most widely deployed FinFET FPGA) running all their DPS resources (7,000 DSP slices) at max performance. **5G Bandwidth** is based on a Massive-MIMO 16x16 radio implementation comparison, comparing Everest to our latest RFSoC devices at 16nm. Everest leverages advanced 5G accelerators for 4X signal processing performance to implement 16x16 800MHz digital radio (which can be deployed as part of larger Massive MIMO antenna arrays). Latest RFSoC devices have signal processing bandwidth (4,000 DSP slices) for 200MHz of 5G spectrum.