Essence of High Speed Connectivity: When Big Data Meets Big Compute

Wendy Chen, APAC IP Sales and Eco-system Group Director
D&R IP SoC DAYS
Shanghai,
Sep 8th, 2018
The Cyber World is Growing BIG and FAST

Facebook
8 Billion video views daily

Amazon
6,000 orders closed each minutes

YouTube
300 hours of video uploaded each minute

NETFLIX
140 Million hours of Content watched daily

WeChat
38 Billion daily messages
Evolution of Data Processing: Mainframe/Client/Cloud/Accelerator

- **Mainframe Era**
- **Client-Server Era**:
  - x86 Speed Wars
  - Multi-Core Wars

- **Server Virtualization Era**

- **Cloud - Software Define and ML / AI Acceleration**

- **Application-Specific Accelerated Cloud**
What Is Happening Today in the Enterprise World?
Per IDC: In 2018, 75% of enterprise developments will include AI / ML / DL

<table>
<thead>
<tr>
<th>Healthcare and Life Sciences</th>
<th>Financial Services</th>
<th>Government</th>
<th>Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Molecular dynamics</td>
<td>• Risk analysis</td>
<td>• Disaster planning</td>
<td>• Smart facilities</td>
</tr>
<tr>
<td>• Drug interactions</td>
<td>• Fraud detection</td>
<td>• Emergency service allocation</td>
<td>• Predictive maintenance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retail</th>
<th>Energy</th>
<th>Transportation</th>
<th>Travel</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Consumer analytics</td>
<td>• Production management</td>
<td>• Driver assistance</td>
<td>• Traveler recognition</td>
</tr>
<tr>
<td>• Product recognition</td>
<td>• Environmental impact</td>
<td>• Autonomous driving</td>
<td>• Intelligent overbooking</td>
</tr>
</tbody>
</table>
Why: Data Is the New Oil!

AI can make the smart grid smarter

AI can make health diagnoses faster, for better treatment and improved health care

AI can make retailers know shoppers pattern, to get them to come back again

AI can make bionic hands know what they are grabbing

20% Boost in energy production

40% Higher chances of finding rare genetic disorders

65% Reduction in lost sales due to logistics and delivery, accuracy in inventory management

10X Faster than contemporary prosthetics
Datacenter Opportunity: ~$18B SAM in 2020
Workload-optimized, high-performance compute, connectivity, accelerators – ML/DL/AI

Hyperscale Optimization CPU
- Workload optimized
- Machine learning
- Deep learning
- Accelerator offloads

Rack-Level Connectivity
- Leaf/spine
- Memory pool (HBM)
- Connectivity / SiP
- Reduced latency
- Mesh / 3D-torus / fabric

Scale Out Clusters
- DNN
- SSD / NVMe
- Coherency
- VM / containers
- Mesh/3D-torus
Market Trend – Datacenter Cloud Connectivity
Majority of connectivity will be 50G/100G by 2020

<table>
<thead>
<tr>
<th>Segments</th>
<th>Speed of Adoption</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1 Cloud Datacenter (&gt;1M Servers)</td>
<td>Early Adopter</td>
<td>10GbE &gt; 40GbE</td>
<td>40GbE &gt; 50GbE</td>
<td>50GbE &gt; 100GbE</td>
<td>100GbE</td>
<td>100GbE+</td>
</tr>
<tr>
<td></td>
<td>Majority Adopter</td>
<td>10GbE</td>
<td>10GbE &gt; 25GbE</td>
<td>25GbE &gt; 50GbE</td>
<td>50GbE &gt; 100GbE</td>
<td>50GbE &gt; 100GbE</td>
</tr>
<tr>
<td>Tier 2/3 Cloud Datacenter</td>
<td>Early Adopter</td>
<td>1GbE &gt; 10 GbE</td>
<td>10GbE &gt; 25GbE</td>
<td>25GbE &gt; 50GbE</td>
<td>50GbE &gt; 100GbE</td>
<td>50GbE &gt; 100GbE</td>
</tr>
<tr>
<td></td>
<td>Majority Adopter</td>
<td>1GbE &gt; 10 GbE</td>
<td>10 GbE</td>
<td>10GbE &gt; 25GbE</td>
<td>25GbE &gt; 50GbE</td>
<td>25GbE &gt; 50GbE</td>
</tr>
<tr>
<td>Enterprise / On Premises</td>
<td>Early Adopter</td>
<td>1GbE &gt; 10 GbE</td>
<td>10GbE &gt; 40GbE</td>
<td>10GbE &gt; 40GbE</td>
<td>10GbE/40GbE &gt; 50GbE</td>
<td>50GbE</td>
</tr>
<tr>
<td></td>
<td>Majority Adopter</td>
<td>1GbE</td>
<td>1GbE &gt; 10GbE</td>
<td>1GbE &gt; 10GbE</td>
<td>10 GbE</td>
<td>10 GbE</td>
</tr>
</tbody>
</table>

Source: Dell'Oro Group 2017, delloro.com
Artificial Intelligence Needs the Most Intelligent Interconnect IP

Intelligent Offload

SoC – Accelerators / Chiplets

Scalability / Efficiency

Higher Connectivity Speeds

Faster Data Processing

Better Data Security

100 Times More Data

10 Times The Speed
What’s the Primary Goal (How): Silicon Design Specific to ML/AI

- Run many jobs in parallel
- Eliminate queue wait times
- Scalable clusters to meet the demand (clients)
- Optimize the resources for appropriate task
- Optimize performance for specific task
Machine Learning Market Landscape

Machine Learning Market Size

- Market Size: $9.1B (2022)
- CAGR: 44%
- Source: Gartner 2018

Machine Learning by Application

- Automotive
- Health Care
- Manufacturing
- Communications
- Education
- Government
- Life Science
- Sports

Types of Machine Learning

- Supervised learning [task driven]
- Unsupervised learning [data driven]
- Reinforcement learning [adaptive]
Deep Learning Semiconductor Forecast

Worldwide Deep Learning Chipset Revenue

Source: Tractica.com
Deep Learning Architecture Example – Intel

Source: https://www.eteknix.com/intel-lake-crest-ai-accelerators/
Drivers for AL/ML/DL Silicon Design

• AI applications are driving the development of new silicon & system architecture
• Process Density & Power key decision driver for silicon design
• Key Trends Influencing AI Silicon Architecture:
  – Processor Architecture
  – Memory Bandwidth
    – Multiple Options: On-Chip, HBM, GDDR
• PCI Express® (PCIe®) Gen 5 / CCIX / GenZ fabric connectivity
• IO Connectivity 25G / 56G / 112G
• Die-Die connectivity within package – scalability
• Silicon Photonics
IP Requirements for Building AI / ML Silicon

- **Connectivity**
  - 25G/32G/56G/112G
  - 32G PCIe® Gen5

- **Memory**
  - HBM
  - GDDR6

- **Inter-die connectivity**
  - D2D
  - 112G XSR
Challenges at 7nm and Beyond
Need for system in package solutions

- New breed of data-intensive vision, graphics, AI, ML, and accelerator applications driving need for increased per socket compute power, increasing monolithic die cost and manufacturability challenges

- Hyper-scale datacenters, driving SoC bandwidth, size and need for hybrid integration (analog, optical, accelerators, etc.)
Die2Die Interconnect
Overview and applications

- Rising monolithic die cost and limitations of on-die integration driving move to die disaggregation and SiP

Typical SiP applications
- CPU to CPU in a multi-core SoC, low-latency coherent interconnect
- DSP arrays, typically to process information from LIDAR
- Switch fabric integration on MCM
- Network ASIC to SerDes PMD on separate die
- Chip to in-package optical engine

Needs:
- High bandwidth
- Low power
- Low latency
- In-package interconnect
112G/56G Landscape – Datacenter, 5G-Mobile, Optics

Optics
- 50G / 100G
- VSR / LR

Mobile – 5G
- 25G / 50G / 100G
- MR / LR
- RRH / BBU

Datacenter
- 25G / 50G / 100G
- MR / LR
- Switch/Fabric/NIC

Every specific segment has its own unique PPA and design requirements—
one for ALL solutions is NOT going to work
Cadence SerDes Advantage

Support
Global support team
250M+ parts in production today
100+ support resources

Time To Market
Accelerating time to market in 7FF node
Proven IP in advanced-process nodes

Characterization / Validation
Extensive validation/QA team
Si tests on multiple parts/corners
HTOL, ESD with stress testing

Feature Rich Design
Multi-protocol PHY
Fractured clocking architecture
Innovative DFE architecture

PPA
Optimized for protocol efficiency
and process nodes

Experienced Design Team
Many years of experience with proven track record of success
Application-Optimized: Broad Portfolio of PHYs

PPA optimized solution

Rate

112 Gbps
- 56G/112G Multi-Protocol Short Reach*
  PAM4

56 Gbps
- D2D Ultra Short Reach*
  NRZ

32 Gbps
- 32/25G Multi-Protocol Long Reach
  NRZ 35dB loss
  PCIe® 5.0, COX-25G
  25/100GBASE-KR

32 Gps
- 10G Multi-Protocol Medium Reach
  USB Type-C, 3.1 Gen2
  PCIe 3.0
  SATA3
  XAUI
  DP 1.4 TX
  SGMII

25/16G Multi-Protocol Long Reach
- CCIX-25G
- PCIe 4.0
- 25GBASE-R
- 10GBASE-KR
- USB3.1 Gen2
- SATA
- RXAU/XAU/XGMII/QSGMII
- XFI/SFI
- JESD204B**
- CPRI®* SRIQ**

5 Gbps
- MIPI® D-PHY™
- USB 2.0

Loss

5 dB 10 dB 15 dB 20 dB 25 dB 30 dB 35 dB

Dedicated PHYs
- USB 2.0
- MIPI-D

Torrent (1-10Gbps)
- Mobile/Consumer
- Low Power
- Small footprint

Sierra (1-25Gbps)
- Mid-End Enterprise
- Multi-Protocol Support
- High Performance
- Low Power

High Sierra (25-32Gbps)
- High-End Enterprise
- High Performance
- Low Power

* In planning
** Some restrictions apply

This slide contains forward-looking statements regarding Cadence's business or products. Actual results may differ materially from the information presented here.

© 2018 Cadence Design Systems, Inc. All rights reserved.
First to Market: Leader in New Nodes
Advanced technology with proven results

- 16FF+ PCIe® PHY in Production!
- 10nm USB PHY in High Volume Production!
- 7nm PHY Silicon Operating at 25/32Gbps
- 16/12nm Si-Proven 16Gbps PCIe 4.0 PHYs
- 16nm 16G PHYs PCIe 3.1 PCI-SIG Compliant!
Time to Market: Integrated PCIe Solution
Controller + PHY with drivers and VIP

- Integration of controller and PHY
  - Time-consuming
  - Risky
  - Adds no value to the product

- Advantages of integrated solution
  - Pre-integrated and verified
  - Reduced risk and cost
  - Accelerated TTM

- Custom driver and documentation
  - Customized for each configuration
Market Leading Solution: PCIe 4.0 / 5.0 and 32G PHY
Live demo at 2018 PCI-SIG Developer Conference (Santa Clara, USA)
Recent Design IP Success

- 2 of 3 Top Mobile Phone Makers Use Cadence PCIe IP
- 1st Commercial PCIe Gen4 and CCIX in 7nm
- DDR5 7nm Functional in Silicon
- 17 out of 25 Top Semiconductor Companies Use Cadence IP
- Automobile Leaders Adopting Portfolio in 7nm and 16FFC