Low Latency Mobile Edge Computing
Devashish Paul
Dir Strategic Marketing IDT
IDT Company Overview

**Founded** | 1980
---|---
**Workforce** | Approximately 1,800 employees
**Headquarters** | San Jose, California

#1 Serial Switching – 100% 4G Infrastructure with RapidIO

#1 Memory Interface – Industry Leader DDR4

#1 Silicon Timing Devices – Broadest Portfolio

800+ Issued and Pending Patents Worldwide

Mixed-signal application-specific solutions
Agenda

- Network Trends
- RapidIO 20-50 Gbps Technology
- Edge Computing Architectures
- OCP Edge Computing Servers and Scale Out
- Open HPAC Lab for Telco Project
The Network is the Data Center

**5G Base Station + Edge Computing Appliance**

- **Ecommerce**
- **Fleet Management**
- **Semi Autonomous Vehicles**
- **Traffic management**

- **Low Latency**
- **Energy Efficient**
- **Analytics Workloads**
- **At Network Edge**

- **RapidIO**
- **IEEE 1588**
- **Timing**
- **RF Products**

- **Memory Interface**
- **Retimers**
- **Sensors**
Ubiquitous Computing:
IDT connects, synchronizes, times and makes sense of the human-and-machine connected world
Network and Data Center Convergence

- Apps moving from data center to co-locate with access node (base station or wired access node)
- Supporting real time communication to mobile devices (phones, cars IOT)
- Tight time synchronization between apps running on distributed servers and in data center
- Need low latency interconnect

Boosting User Experience by Innovating at the Mobile Network Edge

An Introduction to the new ETSI Industry Specification Group (ISG) for Mobile Edge Computing (MEC)

Presented by Dr. Rolf Schuster, Vodafone

Edge Computing an essential element of 5G Rollouts
Network Deployment Architecture is Shifting

Today's WCDMA & LTE with RapidIO®
Cloud Radio Access Network (C-RAN)

Enabling LTE-A → C-RAN Switched Cluster with Processing Nodes
Mobile Edge Computing

Towards 5G/MEC → Co-located CPU and Accelerators
Clustering Fabric Needs

- Lowest Deterministic System Latency
- Scalability
- Peer to Peer / Any Topology
- Embedded Endpoints
- Energy Efficiency
- Cost per performance
- HW Reliability and Determinism

RapidIO Interconnect combines the best attributes of PCIe® and Ethernet in a multi-processor fabric
RapidIO in Edge Computing Appliances

- Heterogeneous compute workloads
- No protocol termination CPU cycles
- Energy efficiency
- 20 to 50 Gbps embedded interconnect
- Mission critical reliability
- Scalable Fat node connect multiple boards in Edge Appliance
- Connect multiple boards at Rack Scale in Central Office or C-RAN
- Push Data Center app use cases into the network

Flexible Solutions
Appliance → Rack Scale
• 10/20/40/50 Gbps per port – 6.25/10/12.5 Gps lane
• 100+ Gbps interconnect in definition
• Embedded RapidIO NIC on processors, DSPs, FPGA and ASICs.
• Hardware termination at PHY layer: 3 layer protocol
• Lowest Latency Interconnect ~ 100 ns
• Inherently scales to large system with 1000’s of nodes

• Over 15 million RapidIO switches shipped
• > 2xEthernet (10GbE)
  Over 110 million 10-20 Gbps ports shipped
• 100% 4G interconnect market share
• 60% 3G, 100% China 3G market share
RapidIO Ecosystem and Market Progression

Performance evolves for multiple system platform generations

- 1st Gen Standard
- 2nd Gen Standard
- 3rd Gen Standard
- 4th Gen Standard

- RapidIO Gen3 (10xN) released with path to 25 Gbaud

- Data Center Computing Storage
IDT Launches Next-Generation RapidIO Switches for 5G Mobile Network Development and Mobile Edge Computing

With over twice the performance used in 4G systems, the low-latency devices exceed the RapidIO 10xN standard and are ideal for 5G, HPC, and Mobile Edge Computing.

IDT and Prodrive Technologies Partner to Develop 100ns Latency, Energy-Efficient RapidIO Switch Appliance Portfolio

Optimized Top-of-Rack switches deliver scalable bandwidth—from 750 Gbps to 4.0 Tbps—for 5G, C-RAN, Mobile Edge and High-Performance Computing, Analytics and Financial Trading.
5G Telco Optimized 100ns 50 Gbps Switch Silicon

- **RXS2448**
  - 600 Gbps Full-Duplex Serial RapidIO® Switch
  - 50 Gbps per port
  - 33 x 33 mm package
  - 48 lanes at 12.5 Gbps
  - Up to 24 Serial RapidIO Ports
  - RapidIO Specification (Rev 3.2) Compliant

- **RXS1632**
  - 400 Gbps Full-Duplex Serial RapidIO Switch
  - 50 Gbps per port
  - 29 x 29 mm package
  - 32 lanes at 12.5 Gbps
  - Up to 16 Serial RapidIO Ports
  - RapidIO Specification (Rev 3.2) Compliant

50 Gbps per port | 300mW per 10 Gbps data | 100ns latency
5G Mobile Infrastructure

**KEY APPLICATIONS**
- LTE-A & 5G baseband unit
- Mobile Edge Computing
- Backplane switching
- C-RAN

**APPLICATION ISSUES SOLVED**
- 50 Gbps per port with 95% link utilization
- 100 ns latency
- Power efficient 300 mW per 10 Gbps

Distributed low latency switching
Optimized for needs of OCP Telco Initiatives
Scalable Low Latency Edge Computing Fat Node

Two types of building blocks
- With native RapidIO end point
- Or with I/O Hub Chip

Accelerators = GPU/FPGA for analytics workloads
CPU for server functions
Co located storage
Can be 19 inch form factor, xTCA or other
Energy efficient computing
Scale out beyond small appliances to rack scale up to 80 nodes per rack (42U) for Central Office deployments

Devashish Paul, Director Strategic Marketing
Edge Social Data Analytics

- Analyze User Impressions on World Cup Final 2014 (Germany/Argentina)
  - HPAC Lab project to analyze World Cup 2014 twitter data using Hadoop and visualize using Tableau public on HPAC Platform
5G Lab Germany: Edge Analytics for Autonomous Vehicles

- Network Edge and In Vehicle Analytics
  - Edge Node Multi processor network
  - GPU/x86/ARM/Open Power based Analytics
  - Low latency RapidIO Fabric
  - In vehicle sensor fusion in real time with low latency
  - Leverage OCP Innovations (Edge Appliance and ToR)
Analytics Platform for the Edge

Launched at Mobile Edge Computing Congress 2015 London UK

Example User Applications
- Analytics
- Deep Learning
- Fabric Management
- Messaging Events
- DMA
- High Performance Sockets
- Device Drivers
- Hardware (NIC/Switch/SoC)

RapidIO|GPU|x86|FPGA|Power|ARM
Low Latency | Energy Efficient | High Bandwidth

Contribute to OCP Telco for Edge Computing Planned 1H 2016
Proposed: OCP Telco Low Latency Switch for Edge Computing Scale Out

- 38 x 20 Gbps ports
- Sub 200W switching power
- Support 42U Rack level scale out
- Available Now

Roadmap to 4.8 Tbps
- 96 x 50 Gbps ports
- Sub 400W switching power
- Supports redundant ports to 42U rack and intra rack scale out
- 2H 2016

5G|Mobile Edge Computing |HPC| Video Analytics | Low Latency Financial Trading
Open High Performance Analytics and Computing Lab

- Low latency scalable RapidIO interconnect to accelerate end market usage
- Key focus areas:
  - Hyperscale Cloud Data Center-Based Analytics
  - High-Performance Computing
  - Autonomous Connected Vehicles
  - Wireless 4G advanced, 5G and Mobile Edge Computing
  - Video Analytics
- Project by project contribution model
- Projects completed or in progress
  - Twitter Analytics on FIFA World Cup Finals
  - Supercomputing at the Edge with GPU
  - CERN LHC target acquisition and data center analytics
  - RapidIO ToR switching
  - 5 G Lab Germany: Mobile Edge Computing for 5G wireless networks with connected vehicles

Accelerate Adoption of OCP Solutions
Interest in establishing OCP Telco lab
Contact: Openhpac@idt.com