Modular Networks
Michael Zimmerman
VP & GM, Networking Group at Marvell
The Problem Space

How heavy is the Great Pyramid?

Biggest Crane can lift \(850\) tons

The Great Pyramid weighs \(5,750,100\) tons

Use **Modularity**

Break the problem to smaller pieces
Why Have We Moved to **Modularity**?

- Standard Tools
- Scale to any size
- Collaboration
- Manageable Building Blocks
- Scale to any size
From One Size Fits All to Scaled Modularity

6.4/12.8Tbps

3.2/6.4Tbps

3.2Tbps Switch

6.4/12.8Tbps
Creating Modular Switching Hierarchy

60% Lower Power, 50% Lower Cost

Any Combination of Ports

Fewer Network Layers

No Software is Needed at the ToR Layer

6.4/12.8Tbps
PIPE – Passive Intelligent Port Extender

Support 802.1BR or any System Header

17mm x 17mm

7w Typical

Line Rate Forwarding Engine

Programmable Packet Header Processing

Integrated CPU for Configuration and Control

6.4/12.8Tbps

4x 25G, 2x 50G, 1x 100G

12x 10G, 3x 40G
Is Fanless ToR Possible?

48x 10GbE + 4x 100GbE

Fanless!
Marvell Is Committed to a Full Portfolio of Port Extenders

**In Production**
- 2x 40GbE
- 2x 25GbE
- 40GbE Port Extender
- 8-24x 10GbE
- 25GbE Port Extender
- 16-24x 10GbE

**Sampling**
- 100GbE
- 4x 25G / 2x 50G
- PIPE
- 12x 10GbE
- 12x 10GbE

**In Development**
- 400GbE
- PIPE2
- 25GbE
Gap in ToR Progression

- 0.64 Tbps (2010)
  - 48x10 GbE
  - 4x40 GbE

- 0.88 Tbps (2012)
  - 48/96x10 GbE
  - 4x40/100 GbE

- 3.2 Tbps (2014)

- 6.4 Tbps (2016)

- 12.8 Tbps (2018)

- 20 Tbps
Gap in ToR Progression

- 2010: 48x10GbE, 4x40GbE
- 2012: 48/96x10GbE, 4x40/100GbE
- 2014: 3.2Tbps
- 2016: 6.4Tbps
- 2018: 12.8Tbps

GAP

0.64Tbps

0.88Tbps

3.2Tbps
Design Optimization for **48x 25GbE + 6x 100GbE ToR**

- 48x 10GbE
- 4x 40GbE
- 48/96x 10GbE
- 4x 40/100GbE

- **1.8Tbps**
  - 48x 25GbE + 6x 100GbE

- **0.88Tbps**
  - 48x 25GbE

- **0.64Tbps**
  - 48x 10GbE

- **2010**

- **2012**

- **2014**
  - **3.2Tbps**

- **2016**
  - **6.4Tbps**

- **2018**
  - **12.8Tbps**
Cost Model: Networking for 100K Servers

**Deployment Cost**

- 24 Server Rack: $1,354,275
- 48 Server Rack: $2,709,200
- 96 Server Rack: $1,875,600

**Power**

- 24 Server Rack: 113 KWatt
- 48 Server Rack: 292 KWatt
- 96 Server Rack: 225 KWatt
Cost Model: Networking for 100K Servers

**Deployment Cost**

- 24 Server Rack: $1,354,275
- 48 Server Rack: $2,709,200
- 96 Server Rack: $4,792,050

- Port Extenders
- Full Switch

- 64% Savings

**Power**

- 24 Server Rack: 113 KWatt
- 48 Server Rack: 102 KWatt
- 96 Server Rack: 454 KWatt

- Port Extenders
- Full Switch

- 57% Savings
Interconnect: The New Modularity Evolution

- **USR**: 2.5D/3D applications
  - 1 cm, no connectors, no packages

- **XSR**: Chip to nearby optics engine
  - 5 cm, no connectors
  - 5-10 dB loss @28 GHz

- **VSR**: Chip-to-module
  - 10 cm, 1 connector
  - 10-20 dB loss @28 GHz

- **MR**: Interfaces for chip to chip and midrange backplane
  - 50 cm, 1 connector
  - 15-25 dB loss @14 GHz
  - 20-50 dB loss @28 GHz

- **LR**: Interface for chip to chip over a backplane
  - 100 cm, 2 connectors
  - 35dB at 14GHz
Interconnect at the Silicon Level

3D Stack
2.5D Chip-to-OE
400G - USR
Ultra short reach

12.8Tbps

Memory

Memory

Memory

Memory

12.8Tbps

400Gbps Ultra Short Reach Serdes
Disaggregating Capacity from I/O at the Silicon Level

Separate the I/O from the Main Die

Separate Serdes Process from Packet Processor Process

Ability to Mix and Match Processes

Ability to Mix and Match FAB Technology

OPEN PLATFORM
12.8Tbps

OCP 2018

STAY TUNED