Evolution from 10 G to 40 G & 100 G

Howard Frazier
IEEE 802.3 HSSG
Geneva, CH
May, 2007
Supporters

- Schelto Van Doorn – Intel
- Ilango Ganga – Intel
- Shimon Muller – Sun
- Andy Bechtolsheim – Sun
- Osamu Ishida – NTT
- Shoukei Kobayashi – NTT
- Paul Kolesar – CommScope
- Bruce Tolley – Solarflare
- David Martin – Nortel
- Piers Dawe – Avago
- Brad Turner – Juniper Networks
- George Oulundsen – OFS
- Robert Lingle, Jr. – OFS
- Ali Ghiasi – Broadcom
- Petar Pepeljugoski - IBM
Introduction

- Multi-port, multi-layer 10 G Ethernet switches are being manufactured in volume today
- Experience with previous operating speeds indicates that higher levels of component integration will be desired in order to reduce cost
Evolution from 10 G to 40 G & 100 G

• 10 G products undergoing cost reduction
  – Single 10 G PHYs ➔ Quad 10 G PHYs
  – Single 10 G optics ➔ QSFP optics
• 10 G port density increasing
• 10 G port cost decreasing
Conceptual 10 G Ethernet switch

Single chip 24 port x 10 G multi-layer switch
Conceptual cost-reduced 10 G Ethernet switch

Single chip
24 port x 10 G multi-layer switch

Power

Quad 10GBASE-KR PHY

Quad 10GBASE-KR PHY

Memory

Control

Quad 10GBASE-R PHY

Quad 10GBASE-R PHY

Quad 10GBASE-R PHY

Quad 10GBASE-R PHY

evolutionary development

QSFP

QSFP

QSFP

QSFP
10 G to 40 G

• Very small incremental effort
  — “shim” between MAC and PHY to make 4 x 10 G behave like a 40 G fat pipe

• Re-use Quad PHYs, QSFP optics
Conceptual 10 G Ethernet switch with 2 x 40 G uplinks

- Power
- Quad 10GBASE-KR PHY
- Quad 10GBASE-KR PHY
- Memory
- Control
- Quad 10GBASE-R PHY
- Quad 10GBASE-R PHY
- Quad 10GBASE-R PHY
- Quad 10GBASE-R PHY

- Backplane or stacking connector
- Quad 10GBASE-R PHY
- Quad 10GBASE-R PHY
- Quad 10GBASE-R PHY
- Quad 10GBASE-R PHY

- 40 G uplink
- Incremental development

- Single chip
  16 port x 10 G
  +
  2 port x 40 G
  multi-layer switch
Conceptual 40 G Ethernet multi-port switch

Single chip 6 port x 40 G multi-layer switch

- Power
- Quad 10GBASE-KR PHY
- Quad 10GBASE-KR PHY
- Memory
- Control
- 40 G uplink
- 40 G Port

Quad 10GBASE-R PHY
Quad 10GBASE-R PHY
Quad 10GBASE-R PHY
Quad 10GBASE-R PHY

QSFP
QSFP
QSFP
QSFP
40 G to 100 G

- Requires new 4 x 25 G PHYs and optics
- 10 x 10 G will be useful for some applications
Conceptual 10 G Ethernet switch with 100 G uplink

Single chip
16 port x 10 G
+ 1 port x 100 G
multi-layer switch
Conceptual 40 G Ethernet switch with 2 x 100 G uplinks

Single chip
16 port x 40 G
+ 2 port x 100 G multi-layer switch

new development
Other applications

• Similar progression can be applied to blade server switches
• Assumes that 40 G host interface controllers will also be developed
  – 40 G is a good match for PCI-e bandwidth
Conceptual 10 G Ethernet blade server switch

Single chip 24 port x 10 G multi-layer switch
Conceptual cost reduced 10 G Ethernet blade server switch
Conceptual 10 G Ethernet blade server switch with 2 x 40 G uplinks

- Single chip
  - 16 port x 10 G
  - 2 port x 40 G
- Multi-layer switch
- Incremental development
- 40 G uplink

Power
Memory
Control
Conceptual 40 G Ethernet blade server switch with 2 x 100 G uplinks

- Single chip
  - 16 port x 40 G
  - 2 port x 100 G
- Multi-layer switch

New development
Summary

• 40 G is an evolutionary step from 10 G
  – small incremental effort
  – significant incremental volume
• 40 G will be the right interface for servers 2010 – 2017
• 100 G development proceeds in parallel
  – Uses a different set of resources from 40 G development