



400G時代を見据えたモジュール/ ケージ/コネクタシステム 「QSFP-DD」

Kazuhiro Kureishi
Cisco Systems G.K.
July 27, 2017

ETHERNET ROADMAP

MEDIA AND MODULES

Ethernet is wired technology and supports a variety of media including bi-directional, twisted pair, coax, multimode fiber and single-mode fiber. Most single-mode Ethernet is the twisted pair or CWDM or CWDM+ with FlexE connectors because close to a billion ports a year are sold.

The graphic below shows multiple types of modules that may be used for copper or optical links. The upper modules are being designed to support ultra-high-speed ports. The xDSP shows the xDSP down to an SP or xDSP. The xDSP enables the lane count to 8. On Board Optics (OBO) enable the highest port counts on switches.

2016 ETHERNET ROADMAP

THE PAST, PRESENT AND FUTURE OF ETHERNET

ETHERNET SPEEDS

© Ethernet Alliance

www.ethernetalliance.org

Developed by Scott Kopp
© Ethernet Alliance 2015. All rights reserved.
03-01

2016 ETHERNET ROADMAP

FLEX ETHERNET (FLEXE)

Defined by the Optical Interconnecting Forum (OIF), FlexE is often called a new generation of Link Aggregation (LAG) and may provide a path to speeds higher than 400G without an IEEE standard. Portions of FlexE claim that higher speeds like 800G are not necessary since these speeds just aggregate many lanes that run at lower speeds. FlexE can aggregate individual Ethernet links and create a FlexE MAC that connects to another FlexE MAC. The illustration below shows a 500G FlexE group that bonds five 100G links together.

The FlexE Group refers to a group of from 1 to n bonded Ethernet Ports. A FlexE Client is an Ethernet flow based on a MAC. Data rate that may not correspond to any Ethernet PHY rates. The FlexE Slice is the layer that maps or demaps the FlexE clients carried over a FlexE group. In the example illustration below three 100G and one 500G FlexE clients connect over the transport network via a FlexE Group that supports 550G/s of bandwidth over 2 wavelengths of 155G/s.

TO TERABIT SPEEDS

ETHERNET INTERFACES AND NOMENCLATURE

Interface	Media	Form Factor	Speed	Standard	Status	Notes
10BASE-T	Twisted Pair	10GBase-T	10 Gb/s	IEEE 802.3-2008	Standard	
100BASE-TX	Twisted Pair	100GBase-TX	100 Gb/s	IEEE 802.3-2014	Standard	
100BASE-SX	Plastic Optical Fiber	100GBase-SX	100 Gb/s	IEEE 802.3-2014	Standard	
25GBASE-SR	Plastic Optical Fiber	25GBase-SR	25 Gb/s	IEEE 802.3-2014	Standard	
40GBASE-SR4	Plastic Optical Fiber	40GBase-SR4	40 Gb/s	IEEE 802.3-2014	Standard	
100GBASE-SR4	Plastic Optical Fiber	100GBase-SR4	100 Gb/s	IEEE 802.3-2014	Standard	
100GBASE-SR8	Plastic Optical Fiber	100GBase-SR8	100 Gb/s	IEEE 802.3-2014	Standard	
200GBASE-SR16	Plastic Optical Fiber	200GBase-SR16	200 Gb/s	IEEE 802.3-2014	Standard	
400GBASE-SR32	Plastic Optical Fiber	400GBase-SR32	400 Gb/s	IEEE 802.3-2014	Standard	
800GBASE-SR64	Plastic Optical Fiber	800GBase-SR64	800 Gb/s	IEEE 802.3-2014	Standard	
1.6Tb/s	Plastic Optical Fiber	1.6Tb/s	1.6 Tb/s	IEEE 802.3-2014	Standard	
3.2Tb/s	Plastic Optical Fiber	3.2Tb/s	3.2 Tb/s	IEEE 802.3-2014	Standard	
6.4Tb/s	Plastic Optical Fiber	6.4Tb/s	6.4 Tb/s	IEEE 802.3-2014	Standard	

Grey Text = IEEE Standard Red Text = In Standardization Green Text = Under consideration in IEEE Blue Text = Non-IEEE standard but complies to IEEE electrical interfaces

To get a PDF version of the roadmap and to find out more about the roadmap, please go to: www.ethernetalliance.org/roadmap/

- ✓ 10Mbpsからスタートして2010年までに6つのEthernet Speedが登場
※10M, 100M, 1G, 10G, 40G, 100Gbps
- ✓ 2016年～2020年頃にかけて新たに6つのEthernet Speedが登場/登場予定
※2.5G, 5G, 25G, 50G, 200G, 400Gbps
- ✓ 今後、ルータやスイッチで200G, 400Gに対応した製品が出荷されるにあたり、どのような物理形状のモジュール/ケージ/コネクタの利用が想定されているのか？

振り返り: トランシーバの歴史(GBIC~SFP:1Gbps)



1998



2002



GBIC

- Switching
- 1RU辺り16Port
- 最大1.5W(1Port辺り)



SCコネクタ

SFP

- Switching
- Routing
- 1RU辺り48Port
- 最大1W(1Port辺り)



LCコネクタ

振り返り: トランシーバの歴史(XENPAK~SFP+:10Gbps)



XENPAK

2003



X2

2004



XFP

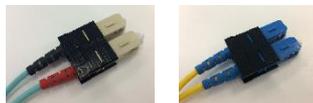
2004



SFP+

2007/8

XENPAK	X2	XFP	SFP+
<ul style="list-style-type: none">• Switching• Routing• XAUI-based• 1RU辺り4Port• 最大9W(1Port辺り)	<ul style="list-style-type: none">• Switching• XAUI-based• 1RU辺り8Port• 最大4W(1Port辺り)	<ul style="list-style-type: none">• Routing• Dual-rate OC-192/10GE• XFI-Based• 最大2.5W/3.5W (1Port辺り)	<ul style="list-style-type: none">• Switching• Routing• 高ポート密度 (1RU辺り48Port)• 低消費電力 (1Port辺り1W/1.5W)



SCコネクタ



LCコネクタ

振り返り: トランシーバの歴史(CFP~QSFP28: 40/100Gbps)



CFP	CFP2	CXP	CPAK	QSFP28
<ul style="list-style-type: none"> • 40G/100G • Switching • Routing • 最大32W(1Port辺り) 	<ul style="list-style-type: none"> • Switching • Routing • 最大6W(1Port辺り) 	<ul style="list-style-type: none"> • Switching • Routing • 最大4W(1Port辺り) 	<ul style="list-style-type: none"> • Switching • Routing • 最大7.5W(1Port辺り) 	<ul style="list-style-type: none"> • 100G(QSFP28) • Switching • Routing • 最大3.5W(1Port辺り)

QSFP+
<ul style="list-style-type: none"> • 40G • Switching • Routing • 最大3.5W(1Port辺り)



SCコネクタ

LCコネクタ

MPO12

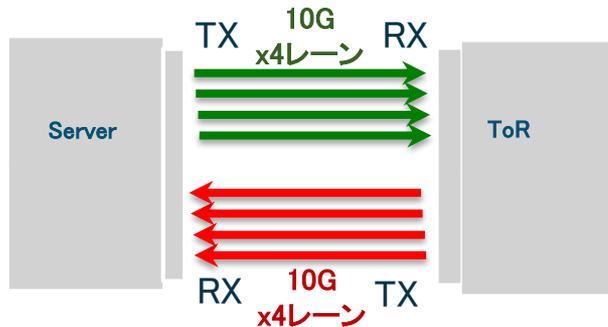
MPO24

振り返ってみて

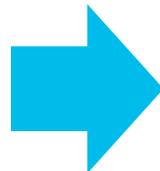
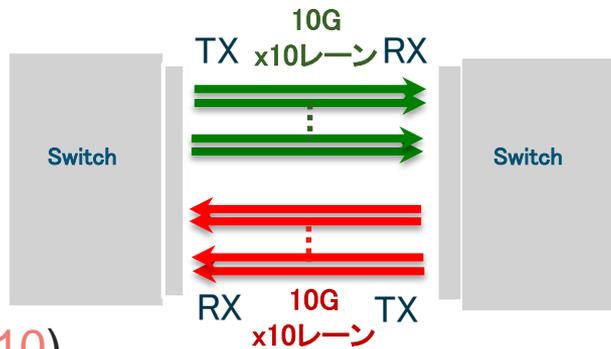
- 基本的に以下の繰り返し
 - Phase1
 - 実現可能なフォームファクター(サイズ、消費電力、EBR等を考慮)で製品化
 - Phase2
 - フォームファクターの小型化
 - 低消費電力化/低発熱へ
 - 1RUあたりのポート密度の向上
 - 低コスト化
 - Phase2'
 - 40G/100G以降については1レーン辺りの増速と必要レーン数の低減

レーン数と1レーンあたりの速度の変化

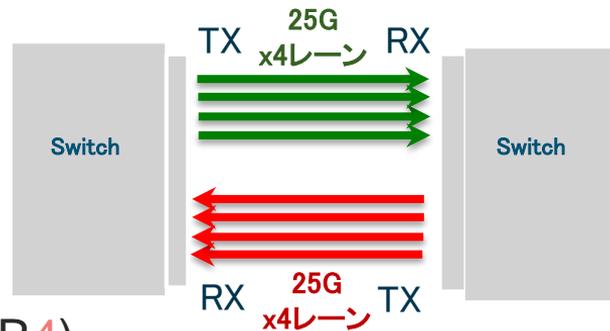
40GbE
(40G-SR4)



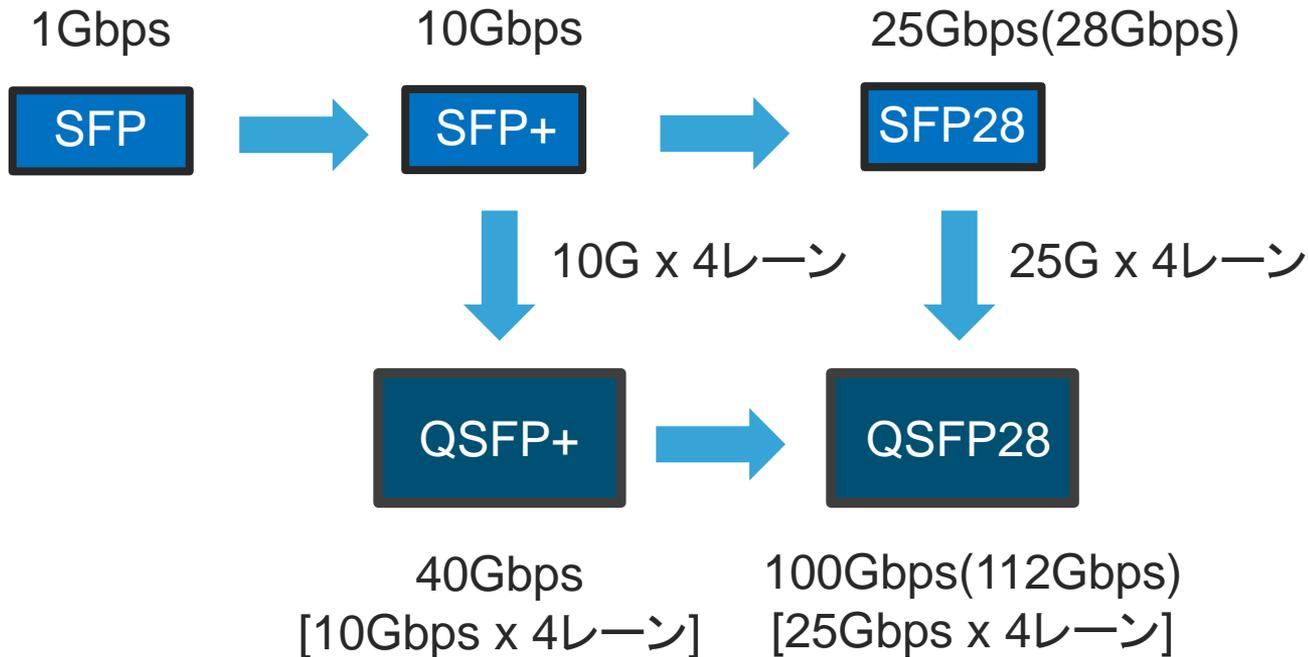
100GbE
(100G-SR10)



100GbE
(100G-SR4)



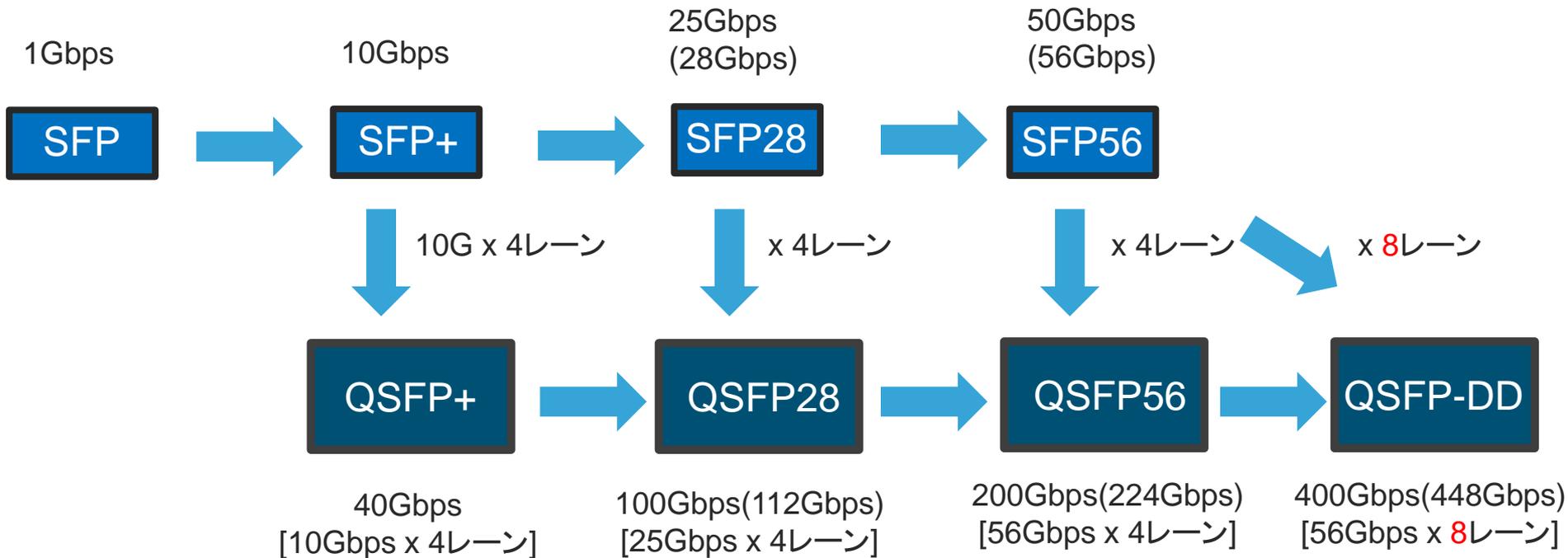
SFP ~ SFP28 / QSFP+ ~ QSFP28



SFP (small form-factor pluggable)

QSFP (Quad small form-factor pluggable)

そして400G対応のQSFP-DDへ

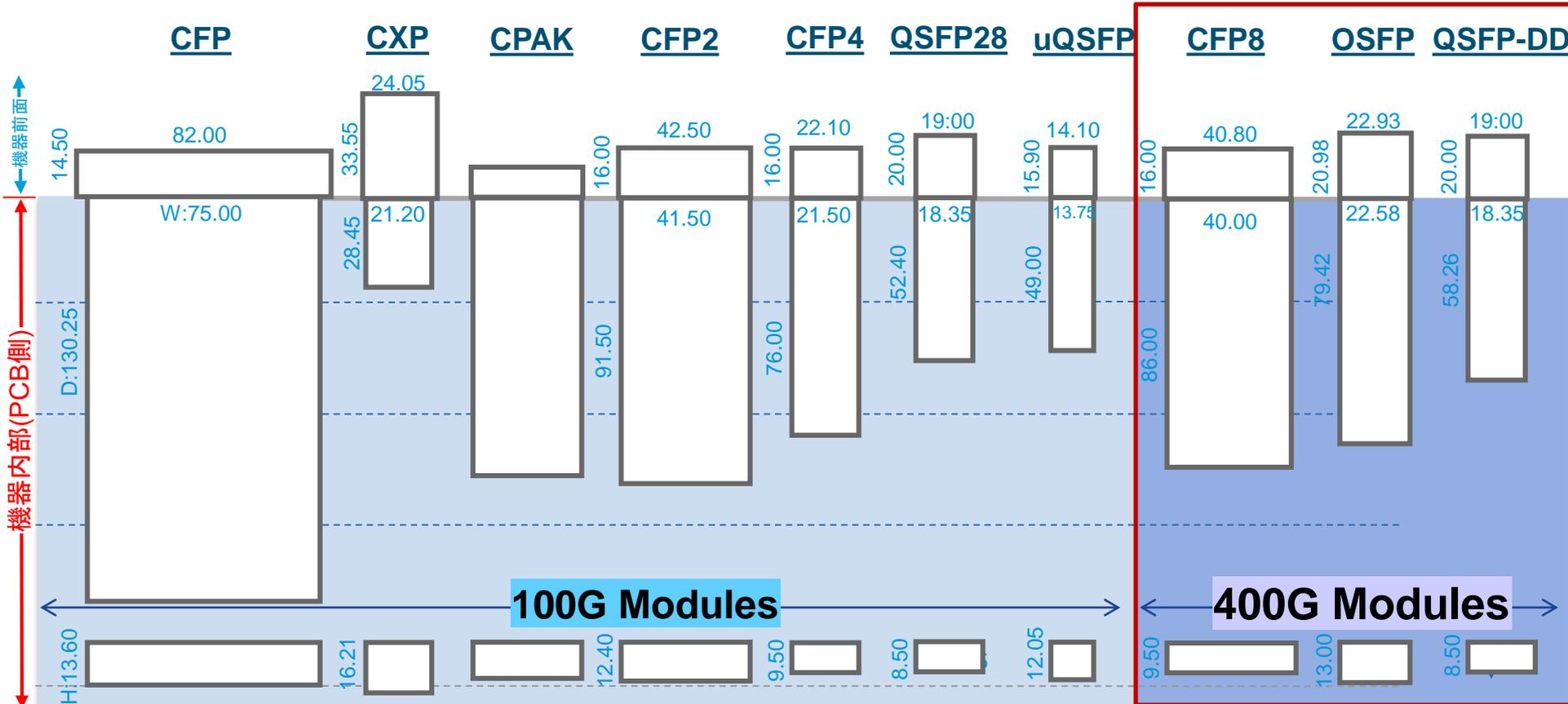


SFP (small form-factor pluggable)

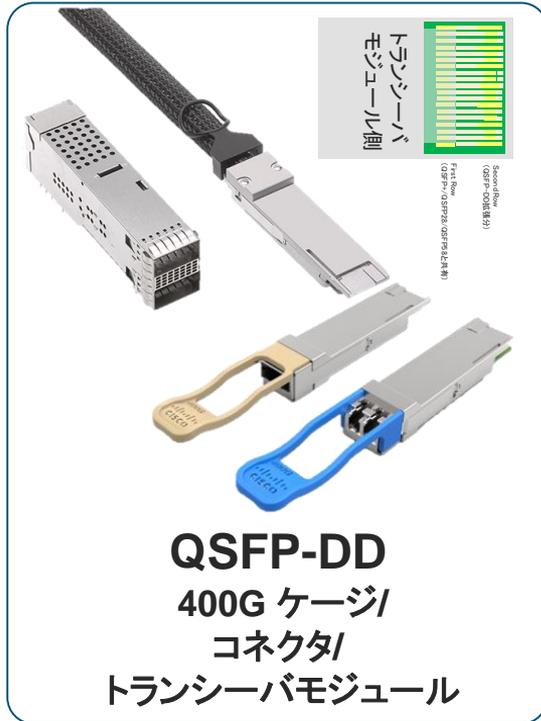
QSFP (Quad small form-factor pluggable)

QSFP-DD (Quad small form-factor pluggable Double Density)

100G/400G トランシーバモジュールサイズの比較



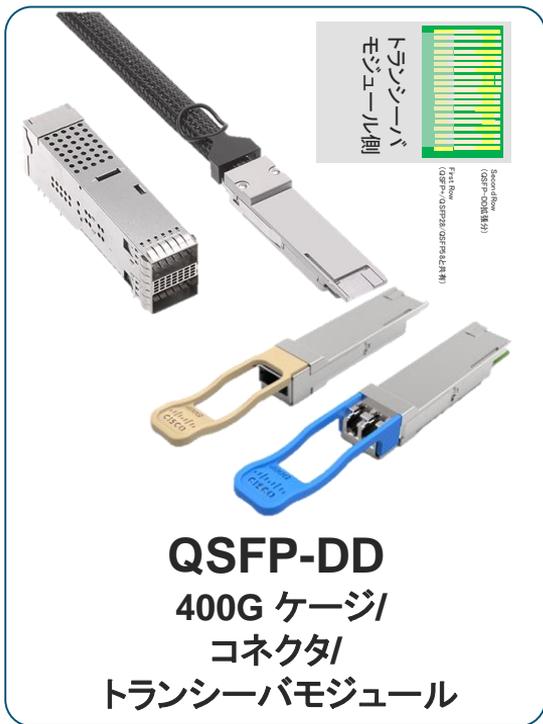
QSFP-DDとは



- QSFP-DD MSA GROUPにはPromoter/Contributorあわせて60社以上が参加
- 高ポート密度に対応可能
 - 1RUで36の着脱可能な400Gポートを実装可能 (理論上、1RUで最大で14.4Tbps)
 - QSFP-DDトランシーバモジュールはQSFP28/QSFP+と同じサイズであり、端子の2列目が増えた分少し長くなっている
 - 8レーン (4レーン x 2列)の50G-PAM4 の電気インターフェースで400Gを実現
 - QSFP-DDケージの物理実装はQSFP+(40G)/QSFP28(100G)/QSFP56(200G)に対して下位互換を提供

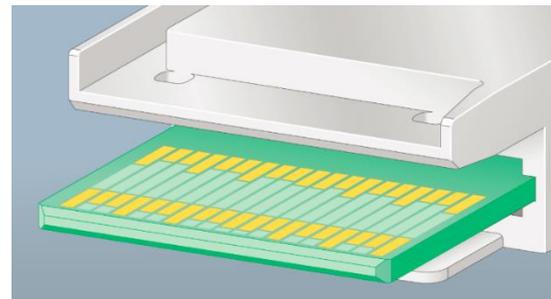
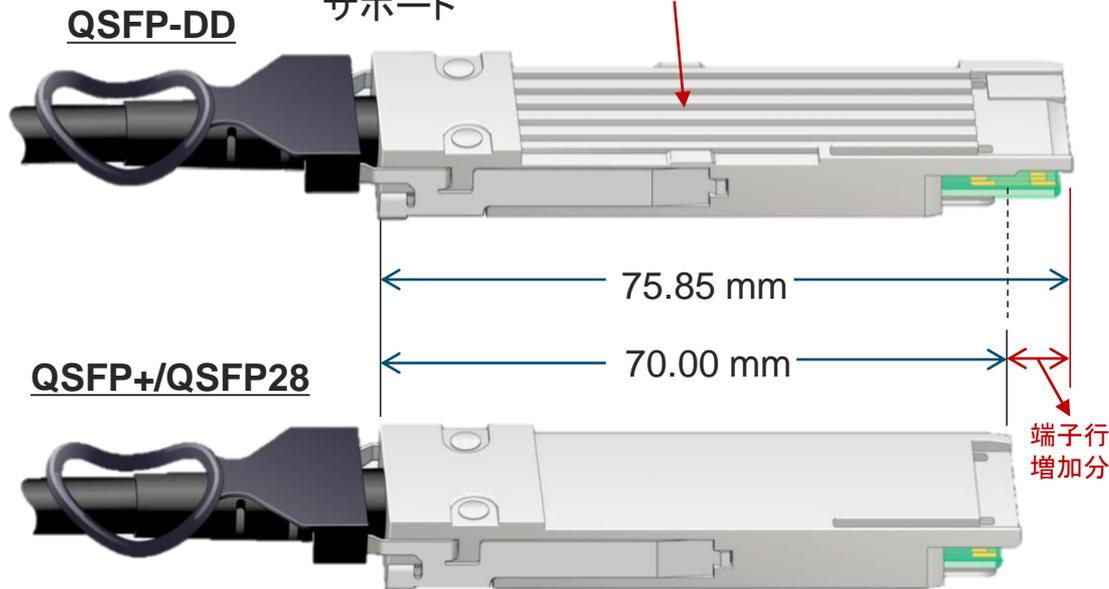
※実際のマルチスピード対応については機器側のPHYやPort ASICのMAC仕様による

QSFP-DDとは



QSFP-DD vs QSFP+/QSFP28

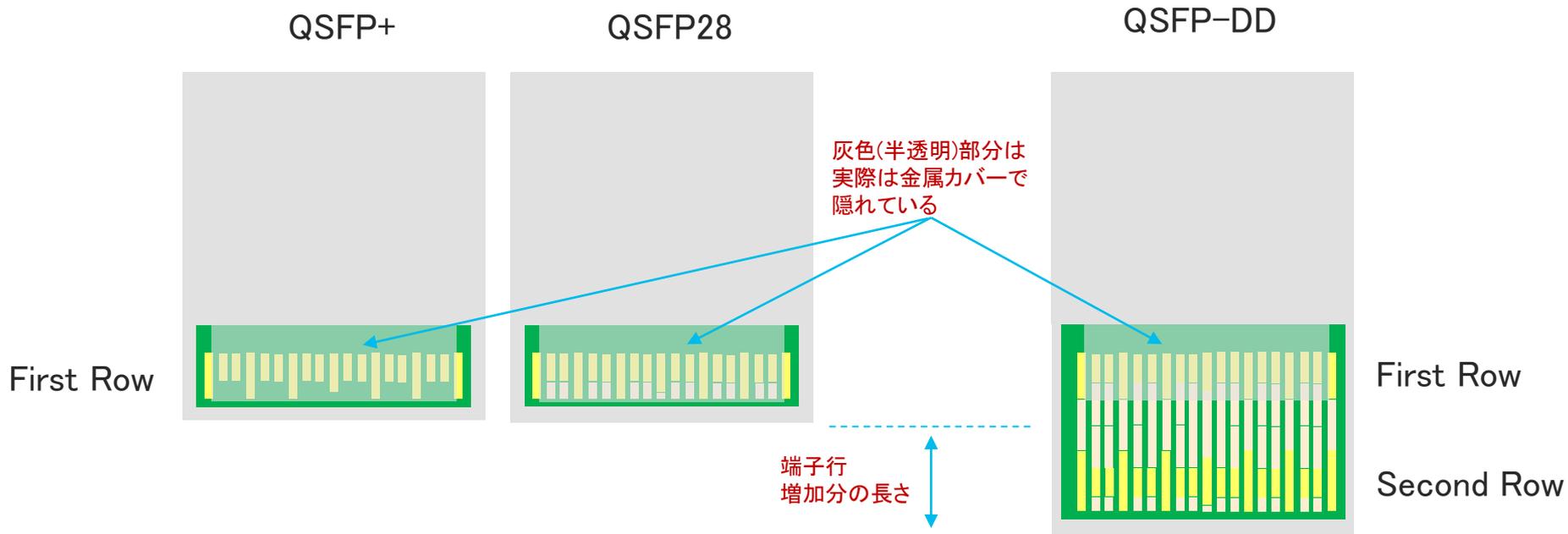
サーマルサポートを改善
※QSFP+/QSFP28の2.5倍以上の消費電力をサポート



基本的にQSFP-DDは
QSFP+/QSFP28と基板上の
端子行が増えた分の長さ
のみが異なる

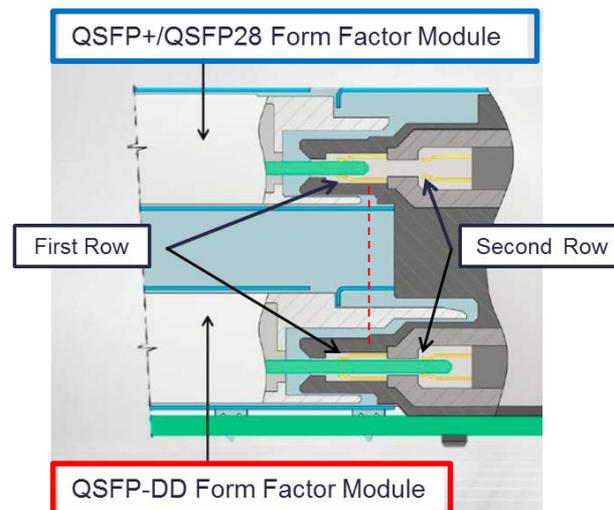
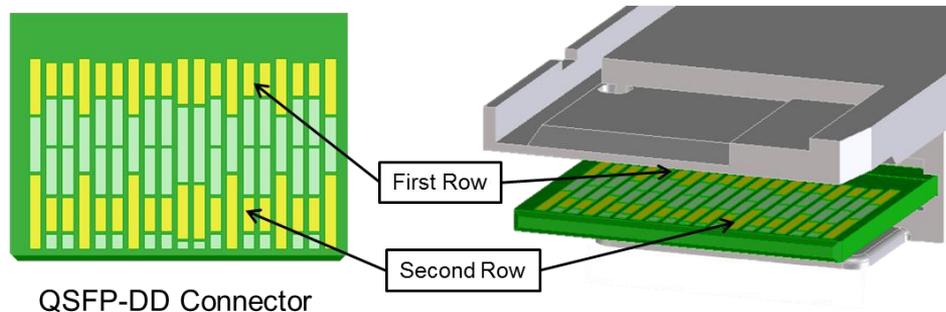
QSFP-DDのケージ(本体側のソケット部分)ではすべてのQSFP (QSFP+, QSFP28, QSFP56)に対して下位互換を提供

QSFP-DDの端子構造と下位互換



QSFP-DDの端子構造と下位互換

- Second RowのPinを追加することで56G-PAM4 x 8レーン = 400Gを実現する
- QSFP-DDポート(ケーシング)は物理レベルでQSFP+ (4x10G), QSFP28 (4x25G), QSFP56 (4x50G-PAM4) の下位互換をサポート
- QSFP-DDのMSAはspecification 2.0として2017年3月に公開 (<http://www.qsfp-dd.com>)



400G時代のフォームファクター: まとめ

- 確認した範囲では400Gに対応可能な主なフォームファクターとしてはCDFP, CFP8, OSFP, QSPF-DDが存在
- 40G/100Gがリリースされた当初と異なり、200/400Gについては最初から比較的小型のフォームファクターでリリースされる事が想定される

