

400Gbpsおよび800Gbpsへの道

Shishio Tsuchiya

shtsuchi@arista.com

質問1

- 400Gインターフェース既に導入していますか？



はい

質問2

- フォームファクターはQSFP-DDですか？

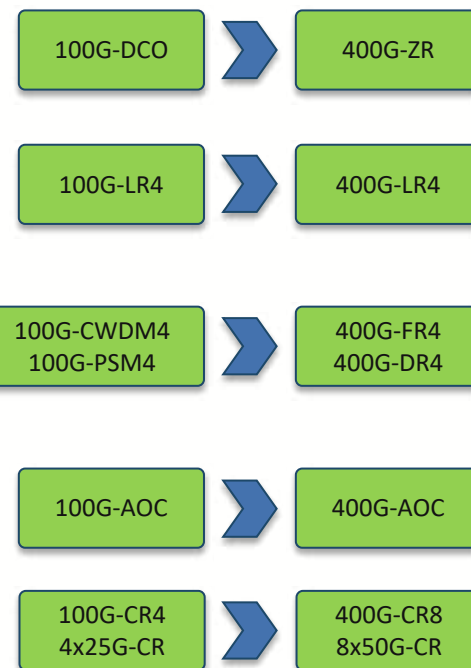
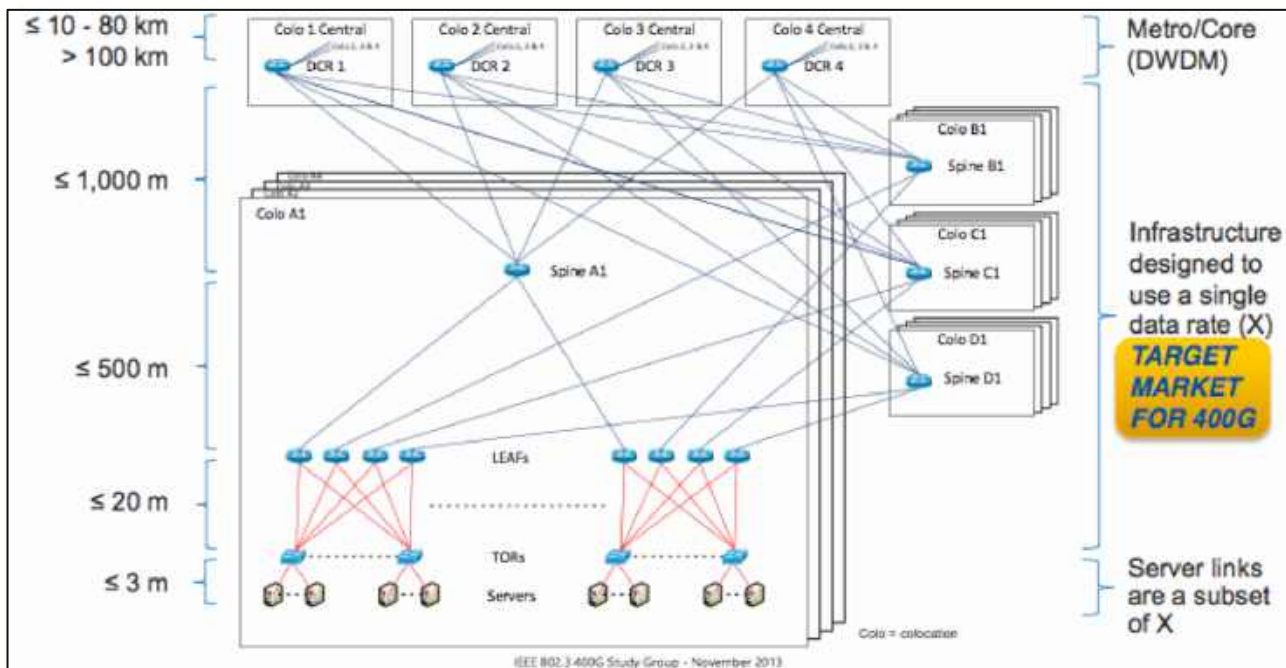


はい

Agenda

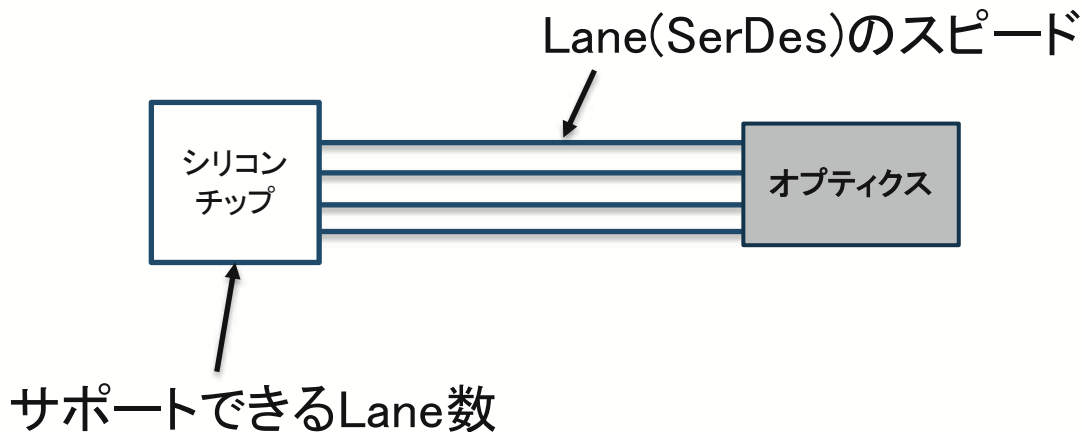
- 400Gbpsロードマップと汎用シリコン
- 400Gbpsフォームファクタ
- 400G ZR / 800 Gbpsへ

クラウドネットワークでの100Gから400Gへの移行



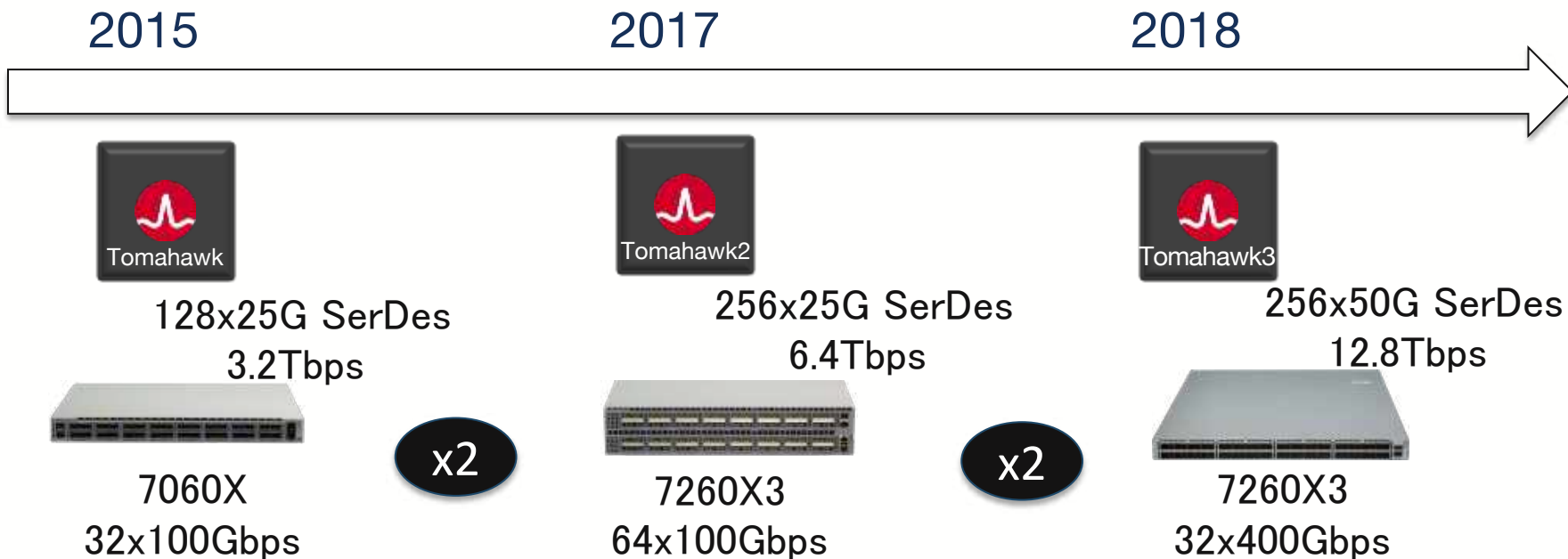
Source: Brad Booth and Tom Issenhuth Microsoft, IEEE 802.3bs 400G

SoC(Silicon On Chip)の容量変化



- SerDesスピードおよび1チップでサポートできるLane数によりSoC容量が決まる
- ムーアの法則(2年間に2倍)をネットワークに

ブロードコム XGS/Tomahawk收容効率の変化



- 收容効率が向上するにつれてフロントポートも増える
- 1年間で2倍の向上が見られた

Lane速度の変化

Lane速度	10Gbps	25Gbps	50Gbps	100Gbps
1X	10G	25G	50G	100G
2X	-	50G	100G	200G
4X	40G	100G	200G	400G
8X	-	-	400G	800G
時期	2011/2012	2015/2016	2018/2019	2020/2021

サーバー
インターフェース

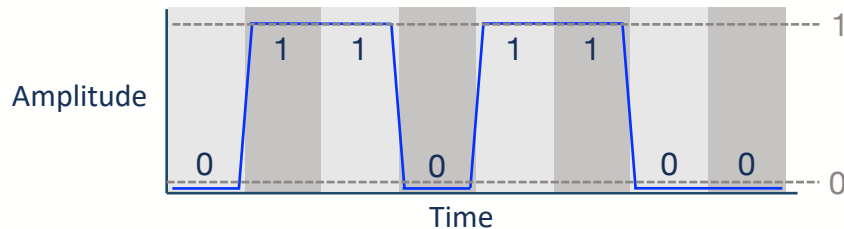
Leaf-Spine
インターフェース



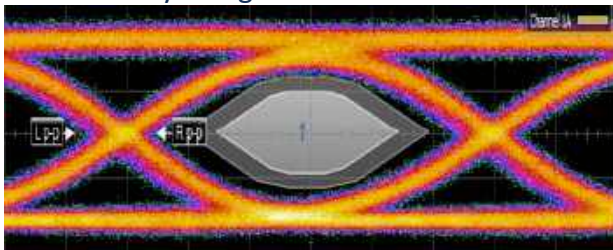
PAM4(Pulse Amplitude Modulation:4パルス振幅変調)

- 400G転送では従来のNRZ(Non Return to Zero)からPAM4に移行
- 01,10,11,00の4つの電圧レベルのパルス信号として伝送

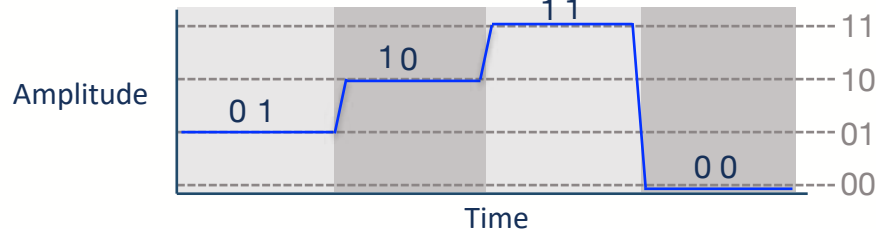
NRZ waveform for data: 0 1 1 0 1 1 0 0



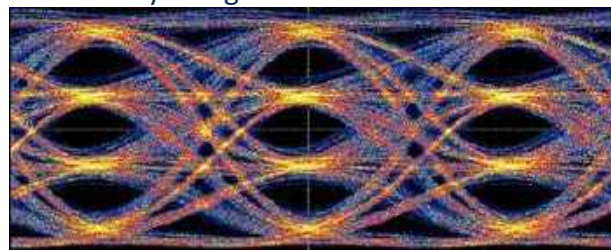
Eye diagram for NRZ data



PAM-4 waveform for data: 0 1 1 0 1 1 0 0

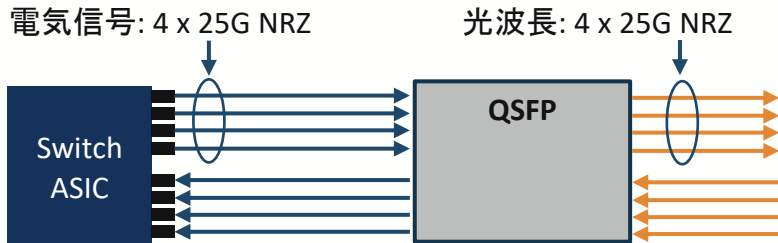


Eye diagram for PAM-4 data



PAM4のインターフェース

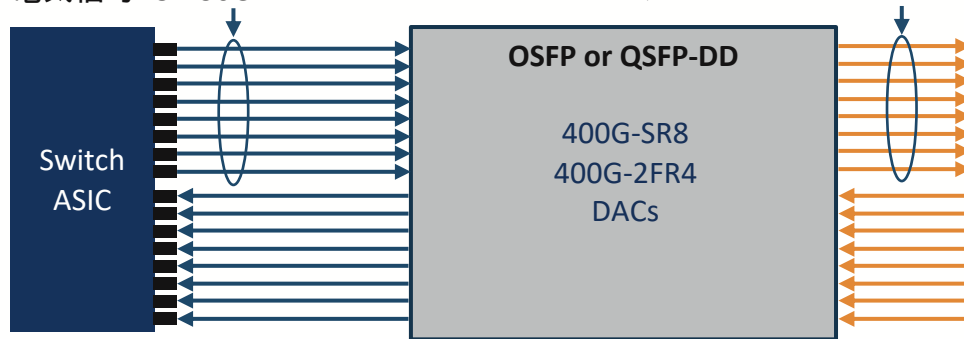
既存100Gオプティクス



400Gオプティクス

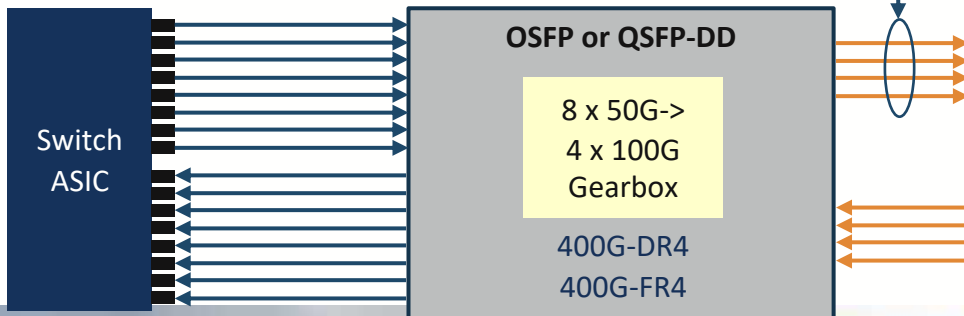
電気信号: 8 x 50G PAM-4

光波長: 8 x 50G PAM-4



または

光波長: 4x 100G PAM-4



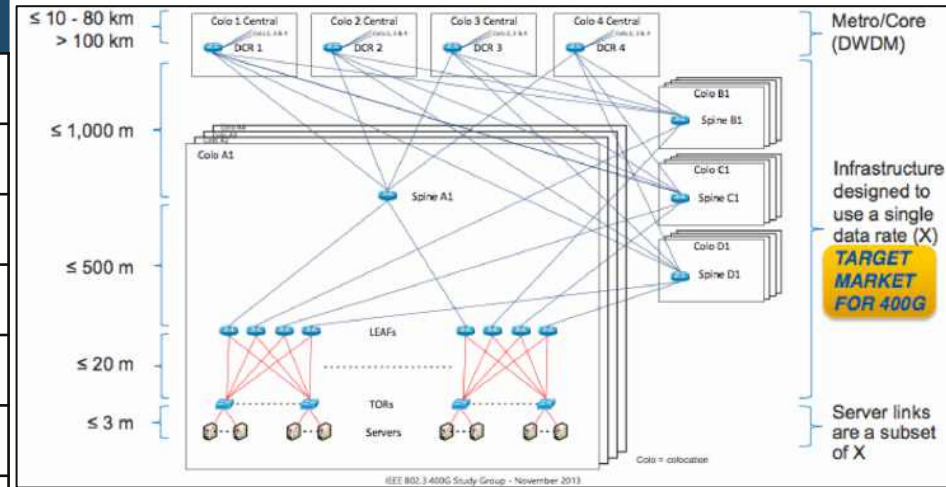
400GE: オプティクスとケーブリング



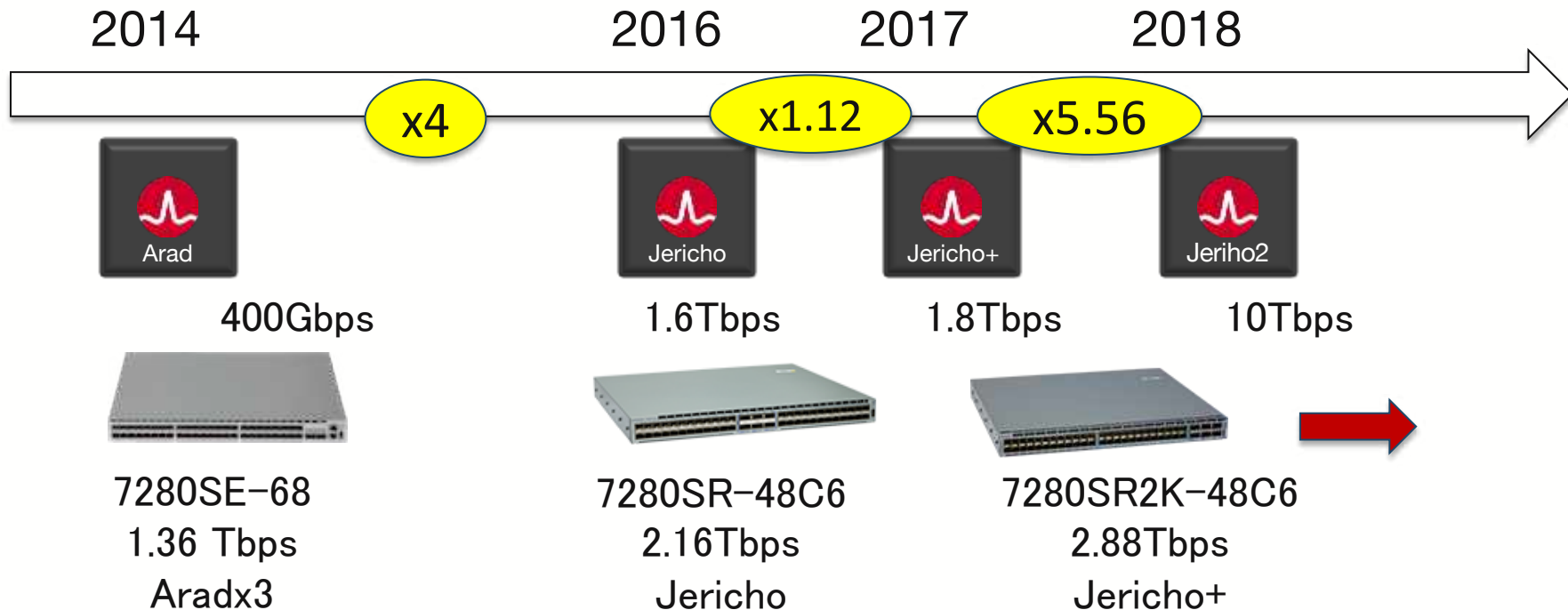
DC Solution

Specification

Intra-rack	3m - Copper	400G-CR8
	30m - Cable	400G-AOC
Intra-DC	100m - pMMF	400G-SR8
	500m - pSMF	400G-DR4
	2km - SMF	400G-FR4
	10km - SMF	400G-LR4
DCI	100km - SMF	400G-ZR

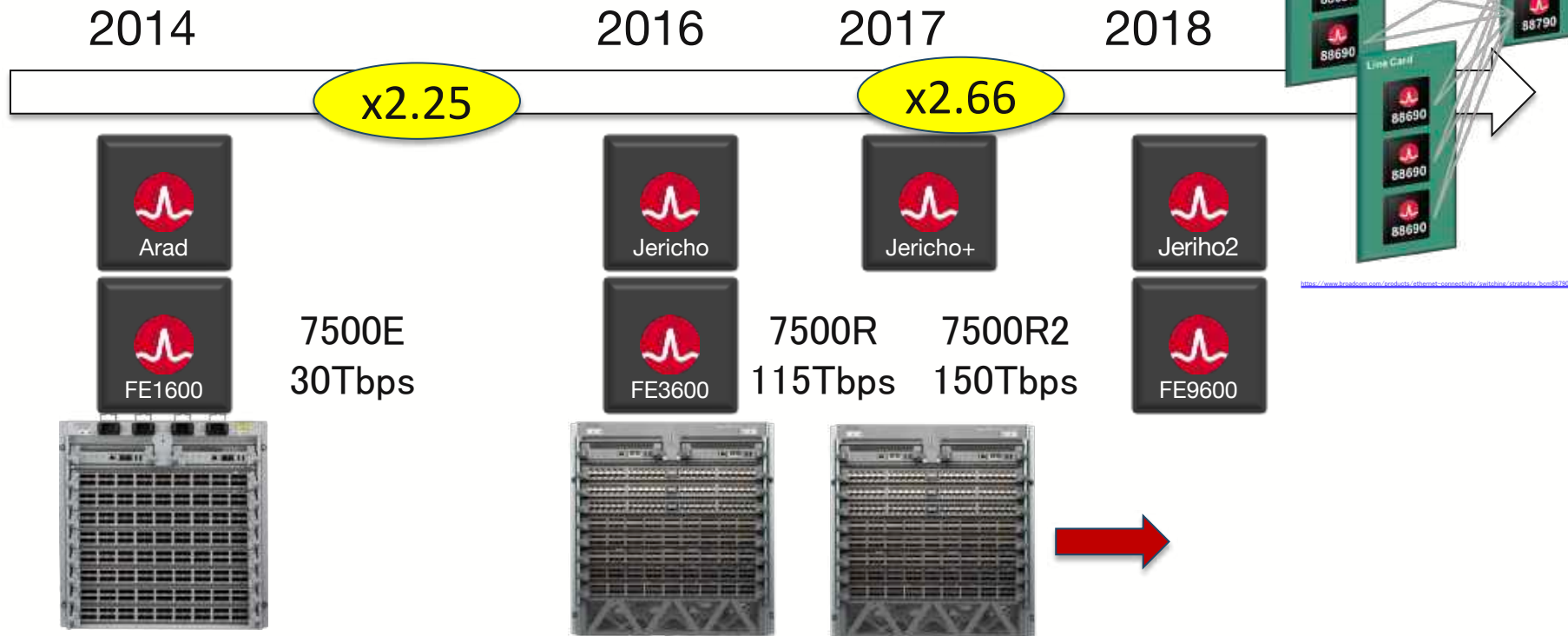


ブロードコム DNX/Jericho収容効率の変化



- VoQ/ラージルーティングテーブル/MPLSの対応のJerichoシリーズはストレージ/サービスプロバイダー領域で多く使用される
- 5年間で20倍強の成長をしている

ブロードコム DNX/Jericho収容効率の変化



- DNXでは同時にFabricチップもリリースされる/Jericho2に対応したFE9600がリリース
- これを用いて、モジュラー型のシステムを構成できる

Agenda

- 400Gbpsロードマップと汎用シリコン
- **400Gbpsフォームファクタ**
- 400G ZR / 800 Gbpsへ

オプティカルトランシーバー

1. 電気信号部分IEEEで定義される
2. SFP/SFP+/QSFPなど
3. MTP/LC/LR/SRなど

3.光学インターフェース
コネクタタイプ

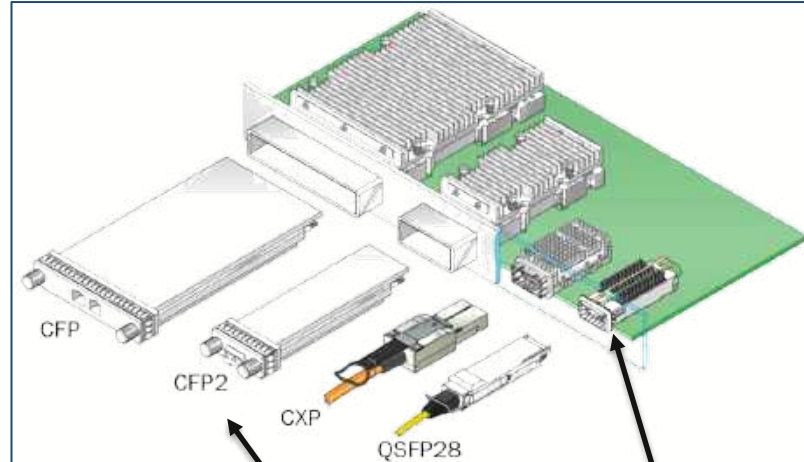


1.電気信号:IEEEで定義

2.メカニカルフォームファクター
アライアンスやMSA(Multi-Source Agreements)で定義

100Gフォームファクター

- CFP MSAがフェーズドアプローチを実施(CFP→CFP2→CFP4)
- SFF MSAがイーサネット/InfiniBand用にQSFP28
- ZRやDCOなどへの対応



ラインカードの対応

2年単位で変わるメディア

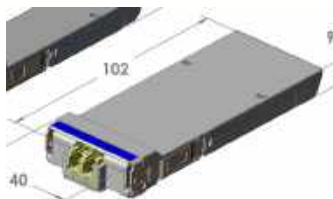
メディア	CFP	CFP2	CFP4	CPAK*	QSFP28
ラインレート	40/100G	40/100G	40/100G	40/100G	4x25-28G
アプリケーション	Ethernet SONET/SDH OTN	Ethernet SONET/SDH OTN	Ethernet SONET/SDH OTN	Ethernet SONET/SDH OTN	Infiniband Ethernet
サイズ	82x145x14	41x104x14	22x92x10	34.8x101.2x11.6	18x52x8.5
最大消費電力	8-32W	3-18W	1.5-6W	4.5-7.5W*	3.5W
時期	2010	2013	2015	2013	2015

Source:
<http://www.fiber-optic-transceiver-module.com/overview-of-100g-client-side-transceivers.html>
 *Cisco Proprietary form factor
http://www.cisco.com/c/en/us/products/collateral/interfaces-modules/transceiver-modules/data_sheet_c78-728110.pdf

100G MSAとは違うアプローチ

- CMOSテクノロジーを持つLightwireを2012年に買収
 - <https://newsroom.cisco.com/press-release-content?articleId=675179>
- 消費電力/サイズで優位を保ちかつ長距離伝送にも活用
 - <https://blogs.cisco.com/sp/cisco-cpak-wins-best-100g-optical-product>
- QSFP28がメインになるまで同一サイズ/ラインカードでの収容が可能

400Gのフォームファクター



	CFP8	QSFP-DD	OSFP
サイズ(HxLxT)	40x102x9.5	18.35x89.4x8.5	22.58x107.8x13.0*
MSA	http://www.cfp-msa.org/	http://www.qsfp-dd.com/	http://osfpmsa.org/
1Uでの最大パフォーマンス	8xCFP8 3.2Tbps	36 x QSFP-DD 14.4Tbps	36 x OSFP 14.4Tbps

- 1U最大14.4Tbpsに収容可能なフォームファクターはOSFPとQSFP-DD

*ヒートシンクを内蔵

OSFPとQSFP-DDの比較



メディア	OSFP	QSPF-DD
800G/ZRに必要な20Wに耐えるか	◎	X
800Gbpsへの上位互換	◎	X
100Gbpsへの下位互換	◎アダプタを使用	◎
DAC最大長	3m	2.5m

OSFPとQSFP-DDを持つ筐体との比較

7060X4での比較

- 32 x 400G OSFP
 - 128 x 100G (パラレルオプティクス&ケーブル)
 - 32 x 100G (OSFP-QSFP アダプタ)
 - QSFP28オプティクスと完全コンパチブル
- 32 x 400G QSFP-DD
 - 128 x 100G (パラレルオプティクス&ケーブル)
 - 32 x 100G (QSFP28オプティクス)
 - QSFP28オプティクスと完全コンパチブル

オプティクス自身にヒートシンクを持つ



熱を放出する為の筐体が必要



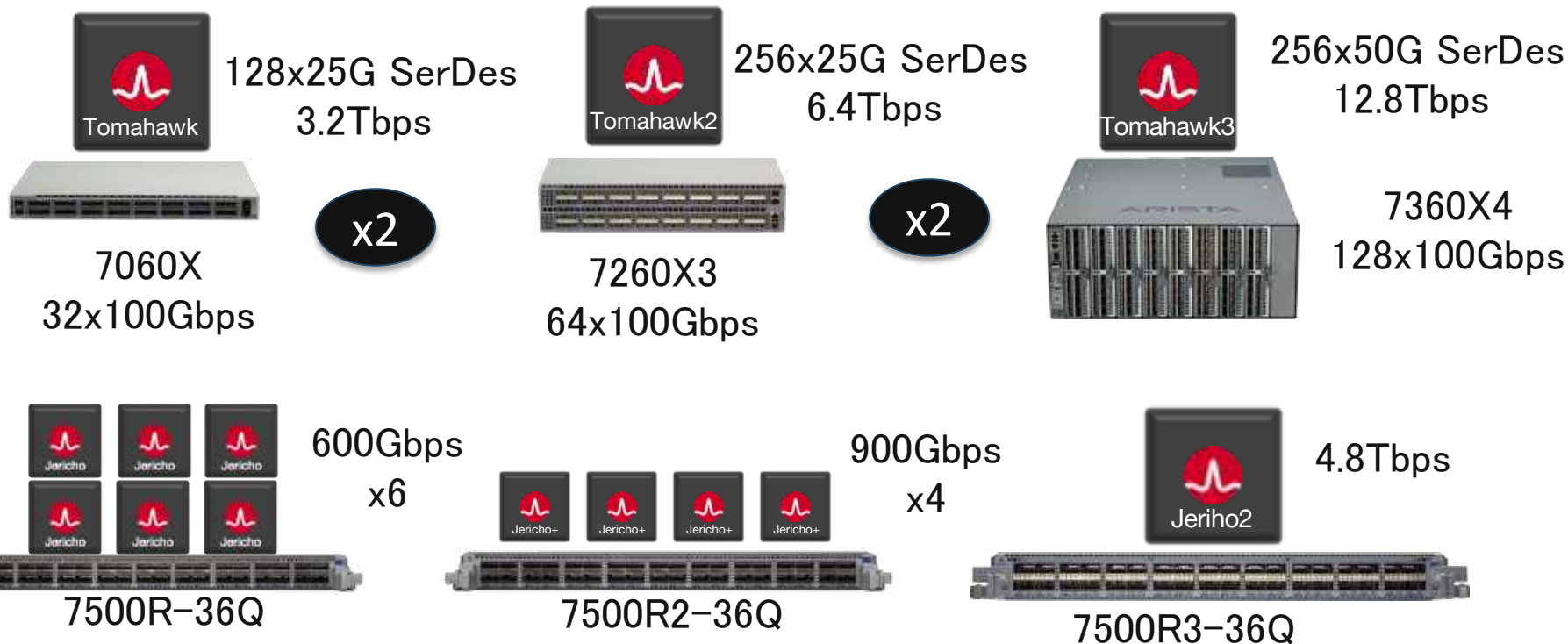
まとめ

- PAM4により1Lane 50Gbpsを実現
- 8x50 PAM4による400Gbpsの実現やGearBoxによる4x100 PAM4を実現可能
- Thomarok3/Jericho2など商用シリコンでのサポートは既に実現済み
- OSFPはヒートシンクが統合されている為、長距離伝送や800Gbpsに対応可能

Agenda

- 400Gbpsロードマップと汎用シリコン
- 400Gbpsフォームファクタ
- **400G ZR / 800 Gbpsへ**

シリコン成長による収容能力の変化



- キャパシティの拡大により大きなシステムの構築
- フロント構成を変えずに、チップ数を減らし、省電力化していく

シリコンスイッチSerDesの予想されるスピードの組み合わせ

年	2017	2018	2019	2020	2021	2022	2023	2024
10G-NRZ	40%	20%	15%	10%	5%	3%	2%	1%
25G-NRZ	60%	80%	65%	40%	20%	12%	8%	4%
50G-PAM4	0%	0%	20%	50%	60%	42%	30%	20%
100G-PAM4	0%	0%	0%	0%	15%	43%	60%	75%

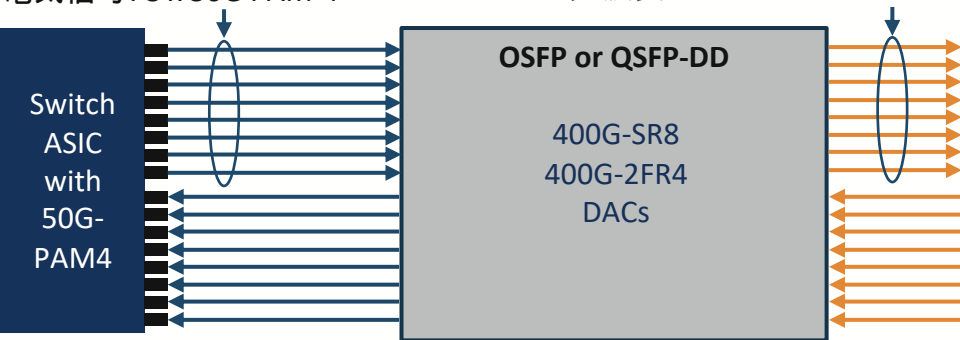
- 50G-PAM4は2019年に増加して、2020年中に25G-NRZを超える
- 100G-PAM4は2021年に増加して、2022年に50G-PAM4を超える

PAM4のインターフェース

400Gオプティクス

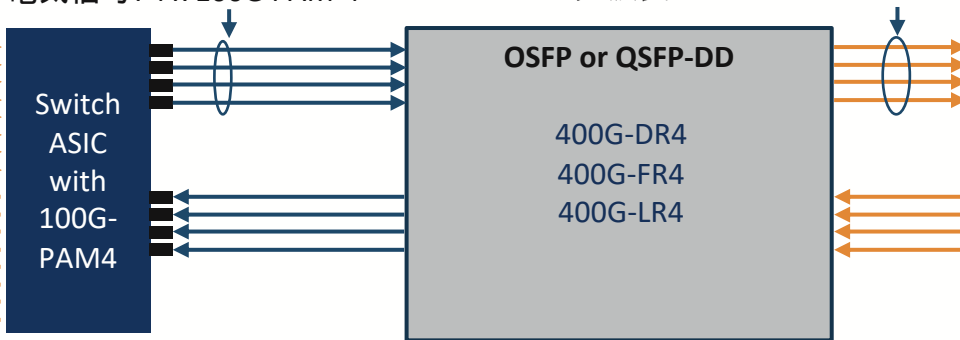
電気信号: 8 x 50G PAM-4

光波長: 8 x 50G PAM-4

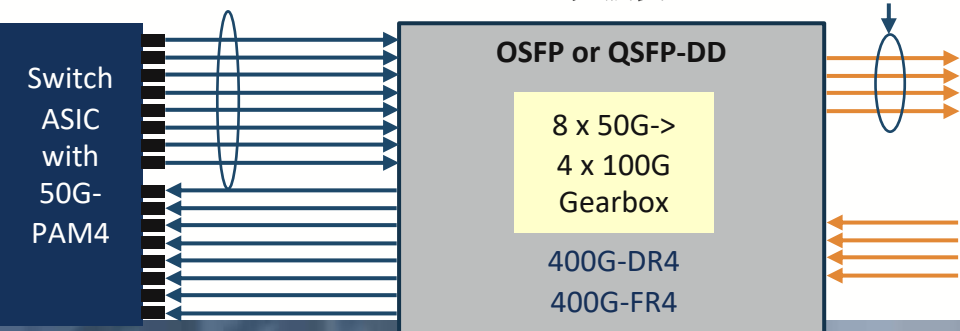


電気信号: 4 x 100G PAM-4

光波長: 4 x 100G PAM-4



または
光波長: 4x 100G PAM-4



100G Lambda
MULTI-SOURCE AGREEMENT

Home · About Us · News · Promoters · Contributors · FAQs · Specifications · Contact Us



<http://100glambda.com/>

IEEEスタンダード/100G Lambda MSA

名称	スタンダード
200GBASE-CR4	IEEE802.3cd
400GBASE-SR8	IEEE802.3cm Breakout 8x 50G-SRはIEEE802.3bs
400GBASE-DR4	IEEE802.3bs Breakout 100G-DR(IEEE802.3cd)
400GBASE-FR4	100G Lambda MSA http://100glambda.com/
200GBASE-FR4	2FR4は 2x 200GBASE-FR4
100GBASE-DR	IEEE802.3cd
100GBASE-FR	100G Lambda MSA http://100glambda.com/

400G-ZRおよび400G-ZR+

- 業界最初のマルチベンダーDWDMスタンダード
- デジタルコヒーレントオプティクスによる着脱可能
- ZR+:400G/300G/200G/100Gスピード対応
 - 400G-16QAM
 - 300G-8QAM
 - 200G-QPSK
 - 100G-QPSK
- 400G-ZR :100km/15W
- 400G-ZR+: 1000km/ 20W

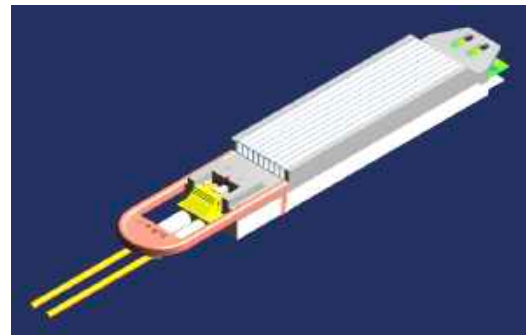
40ポートSFP+
2ポートCFP2-DCO



48ポートSFP+
6ポートQSFP

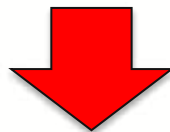
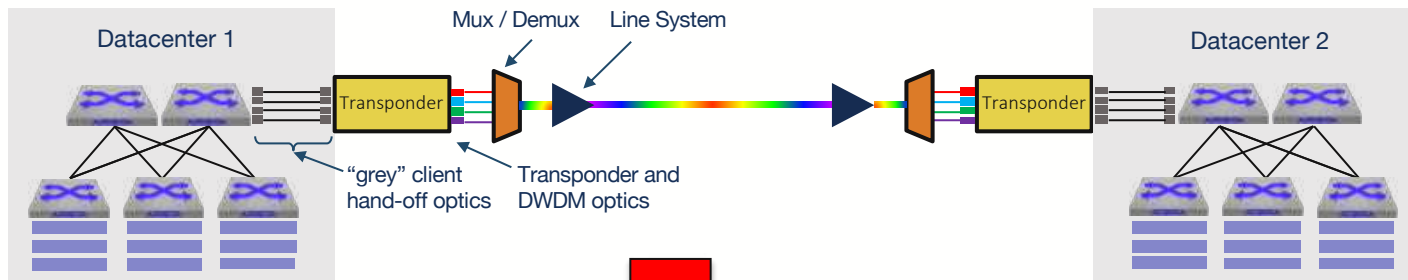


32ポートQSFP
4ポート OSFP



OSFPの特徴を生かしたソリューション

<https://www.arista.com/jp/company/news/press-release/10127-pr-20200309>



チューナブルな400G-ZR



カラーレスな
Mux/Demuxファイバー

OSFP-LSモジュール



Line クライアント
コネクタ コネクタ

光増幅器
(EDFA)

ARISTA

3.2Tbps 120kmのDWDM伝送を1つのボックスで



- 3.2Tbps(400x8チャンネル)の帯域をデータセンター間(120kmまで)をOSFP オプティクスで実現

800Gモジュール

名称	仕様	想定電力
Dual 400G-FR4	8x112G-PAM4電気インタフェース Dual CS Duplexファイバー	18-20W
800G-FR8/LR8/CWDM8	8x112G-PAM4電気インターフェース Dual LCファイバー	18-20W
800G-OSFP-ZR(100km)	800G-16QAM 8x100G SerDes電気インターフェース ファイバー毎に52Tbps(C+L Band)	20W



- OCP2018などで112G-PAM4のデモンストレーションを実施
- 800G/Dual 400Gでは20Wに耐えうるフォームファクターが必要
- OSFPはDACから800G-ZRすべてのシナリオでサポートが可能

まとめ100GSerDes時期予想

- 2020年:最初の100G SERDESスイッチチップとオプティクスモジュールが到着
- 2021年:最初の100G SERDESシステムの製品が出荷
- 2022年:100G SERDESが主要になる

QSFP-DDの800Gbpsおよび400G-ZRへの対応

- Cisco Announces Intent to Acquire Acacia Communications, Inc.
 - <https://www.cisco.com/c/en/us/about/corporate-strategy-office/acquisitions/acacia.html>
- EIGHT INDUSTRY LEADERS COLLABORATE TO INCREASE PLUGGABLE TRANSCEIVER SPEEDS TO 800 GBPS
 - <http://www.qsfp-dd800.com/index.php/2019/09/17/eight-industry-leaders-collaborate-to-increase-pluggable-transceiver-speeds-to-800-gbps/>
- 400ZR: Accessible 400G for Edge DCIs and Beyond
 - <https://acacia-inc.com/acacia-resources/400zr-accessible-400g-for-edge-dcis-and-beyond/>

まとめ

- 下記をまとめた
 - 400Gbpsのフォームファクター
 - 商用シリコンの対応状況
 - 400G-ZRへの対応および800Gbpsへのロードマップ



Thank You

www.arista.com