

400G時代の相互接続 ～ 100G-LR1 って検討されていますか？ ～

BBIX株式会社

堤 友理
空谷 大成

- **堤 友理(つつみ ゆり)**

- 2017～2022
 - NTT Com 株式会社 新卒入社
 - IMF 出向、JPNAP サービスの開発～設計構築～SO～運用保守
 - NTT Com 帰任後、SDPF 関連のサービス企画
- 2023～
 - ソフトバンク株式会社 転職
 - BBIX 出向、BBIX サービスの設計構築
 - タイ拠点展開、伝送装置や光SWの検証/導入、国内NW増強 etc.



▲ タイ深夜作業後にDC前にいたネコちゃん

- **空谷 大成(もくたに たいせい)**

- 2019～2021
 - BBIX 株式会社 新卒入社
 - RPX 展開現地構築
 - 各新規拠点 機器選定 調達 設計 現地作業
- 2022～
 - 主体を新規設計検証部隊へ移動
 - 不具合再現・検証
 - 次世代スイッチ導入を着々と・・・

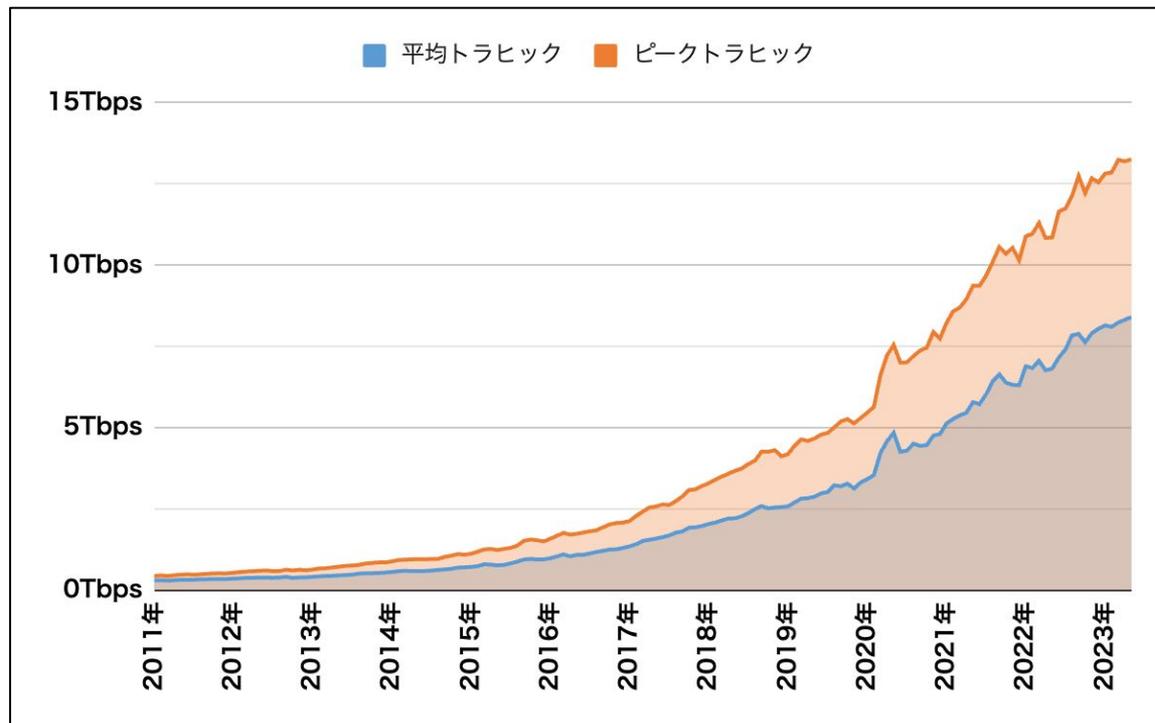


Agenda

- **トラフィック動向**
 - 続く急速なトラフィック増加による課題 / 現場の苦勞
- **400G I/F 導入**
- **400G 時代における 100G 接続**
- **100G-LR1 検討**
 - 100G-LR1 とは
 - 100G-LR4 との比較
 - 期待するメリット
 - ちょこつと検証
- **サマリ**
- **議論したいこと**

加速度的に伸びるトラフィック

- 日本国内主要IXにおけるトラフィックの集計
 - 2011年1月～2023年5月

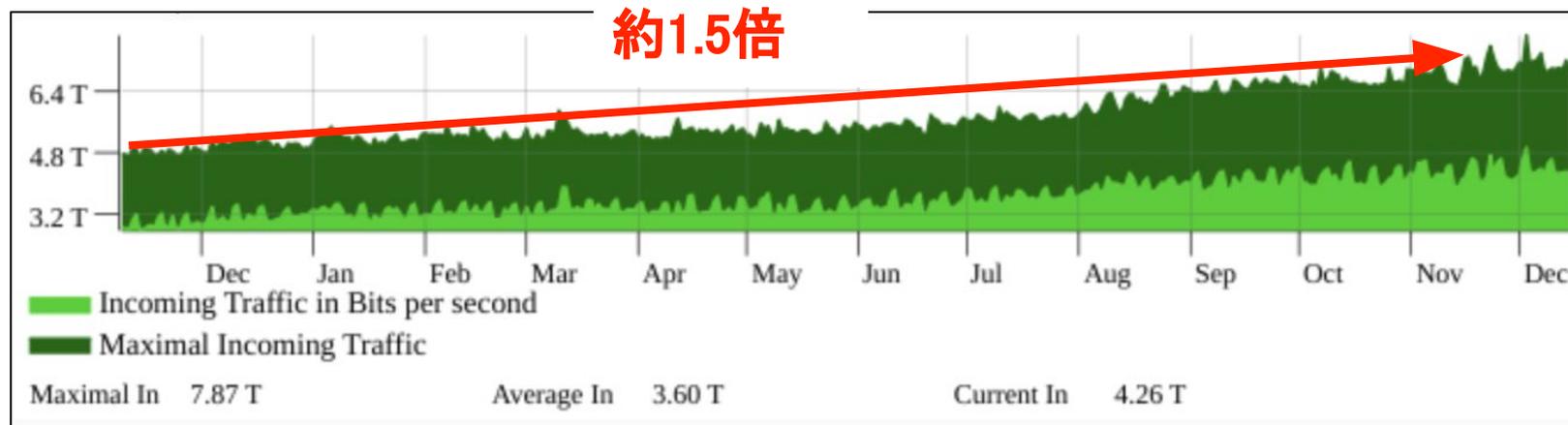


**2020年に上昇加速
直近1年で1.3倍**

- 直近1年間の BBIX トラフィック

- 1年間でピークのトラフィックが 1.5倍になっている
- 2023年において既に1.5倍が見える
- 今後さらなるトラフィック増加によりバックボーンの増強が必要になる

- 現場は輻輳させないようにピーク合わせ対応が必要...



- 現場の苦勞例

- **トラフィックに耐える機器の大量調達**

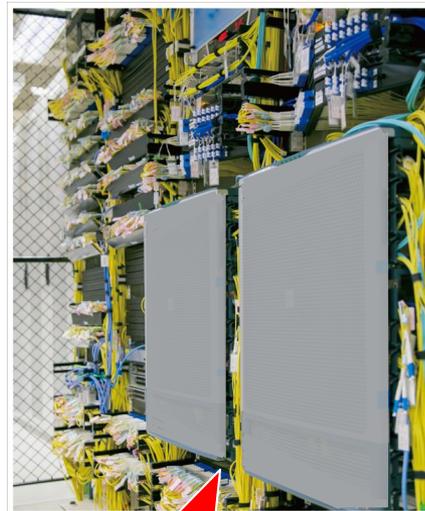
- 納期がかかるので出来るだけ購入しておいた方が良いが、予算が・・・
- トラフィック急増、輻輳までに間に合うか勝負。。ヒヤヒヤする。

- **大量ケーブル調達および配線**

- 必要以上のケーブル購入は企業によっては予算的ダメージが。。
- 配線が増えすぎて、ラックに触ったら事故が発生しそうで怖い。

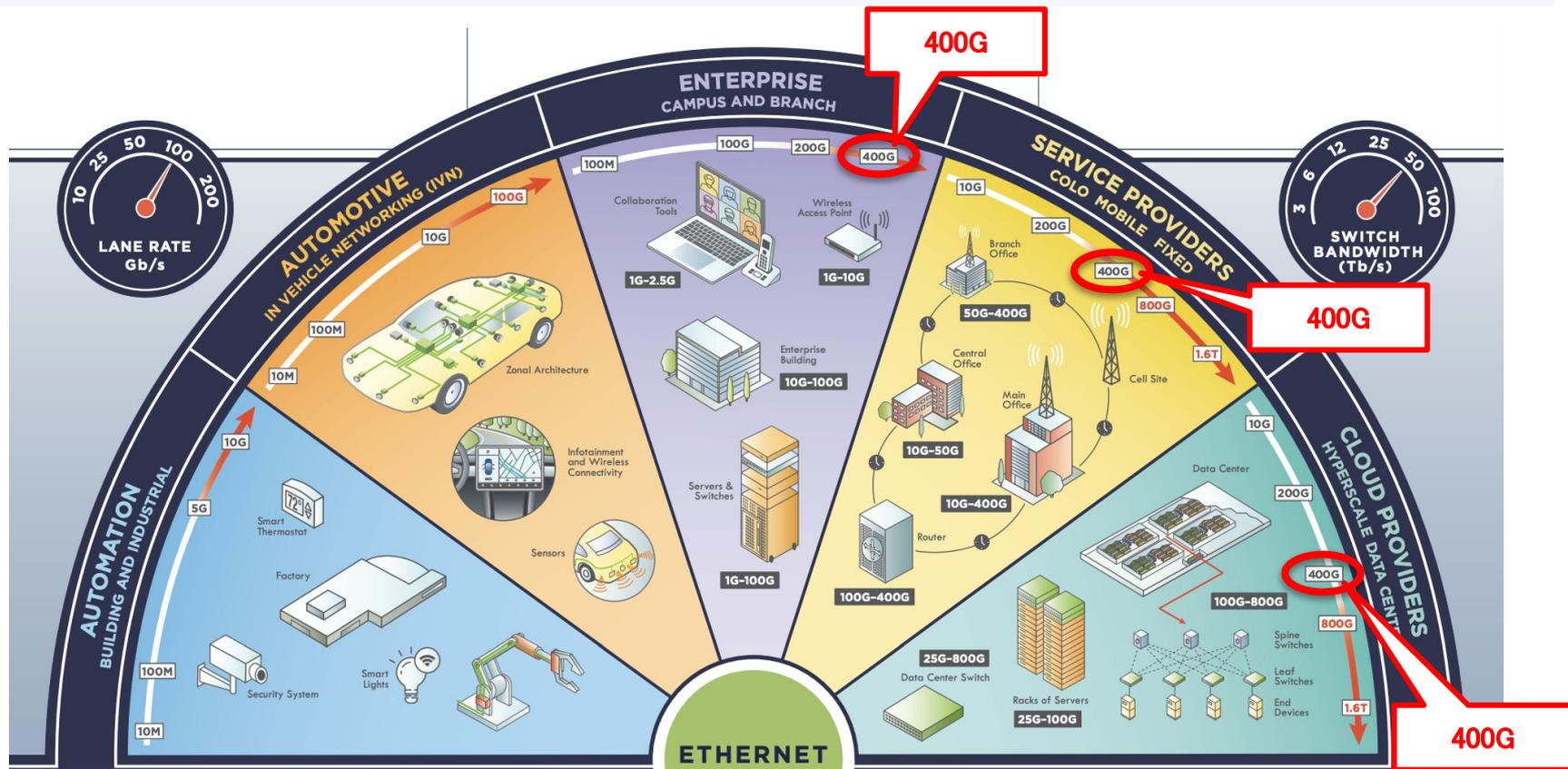
- **機器ラッキングスペースの確保**

- 配線も機器もいっぱい追加機器置けない
- ラック代、電気代高騰で新規契約が。。



**現場ピンチ
画期的な解決策
必須**

400G Interface の登場



Ethernet Alliance, EthernetRoadmap-2023-Website-REV-March-17,
<https://ethernetalliance.org/wp-content/uploads/2023/03/EthernetRoadmap-2023-Website-REV-March-17.pdf>

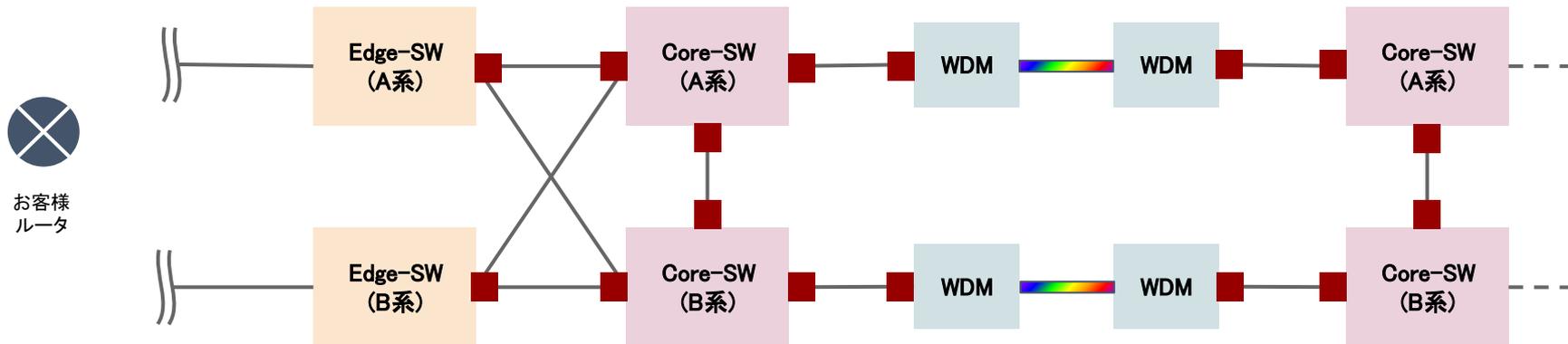
400G Interface の登場

- 各社より 400G I/F 対応機器が登場
 - CISCO
 - Nexus 9364D-GX2A
 - JUNIPER
 - PTX10001-36MR
 - Arista
 - 7388X5
 - HPE
 - Networking Comware 5960
 - Ruijie
 - RG-S6580-48CQ8QC
 - QuantaMesh
 - BMS T9032-IX9
 - Edgecore
 - AS9716-32D
 - Nokia
 - 7220 IXR-D5

...etc.

400G Interface の導入

- 課題解消に向けて、BBIX でも 400G I/F を導入予定
 - ■の区間に400G I/Fを順次導入していく



収容効率向上

ラックスペースや電力の利用効率向上



機器の急速なポート枯渇抑制

トラフィック急増による緊急調達の納期問題解消



ケーブルコスト削減、逼迫緩和

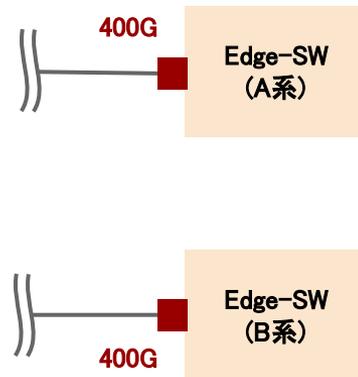
400G Interfaceの需要

- 課題 / 現場の苦勞 はどこの現場も同じはず
- 大手コンテンツ事業者を中心に、IX接続ポートも 400GbE の需要の声あり
 - → BBIXでも提供品目として 400GbE をリリース予定(2024/04～)

- 収容効率向上**
ラックスペースや電力の利用効率向上
- 機器の急速なポート枯渇抑制**
トラフィック急増による緊急調達の納期問題解消
- ケーブルコスト削減、逼迫緩和**



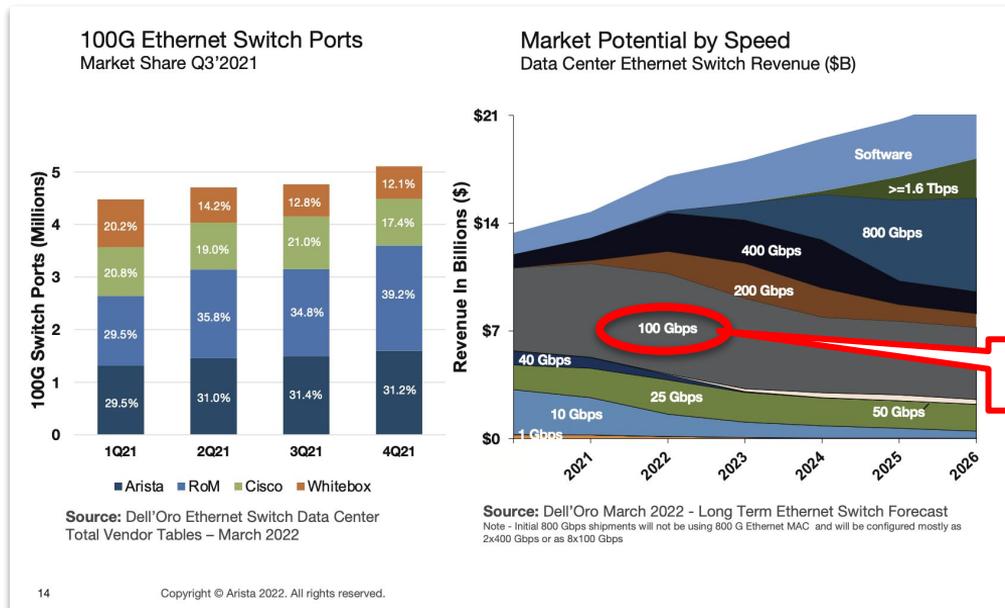
お客様
ルータ



→ 段々と各事業者で400G I/F の導入が進んできている

現在の主力は 100G Interface

- とはいえ、実際にはまだ 100G I/F が主力ではないでしょうか
 - BBIXのバックボーンではまだ全て 100G I/F を束ねている箇所もある
 - お客様向けのIX接続ポート提供品目も、2024/01 時点では最大で 100GbE



→ そこで...

QSFP28 Single-lambda 製品

- 400G 光モジュールを B/O して使うこともできる 100G Single-lambda 製品
- その中でも、現在 DCI 接続において主流な 100G-LR4 を補完する **100G-LR1** に注目
 - → BBIX でも 100G-LR1 の申込受付開始 (2024/01~)

伝送距離	500m	2km	10km	40km
ファイバ種別	SMF	SMF	SMF	SMF
モジュール規格	100G-PSM4 100G-DR1	100G-CWDM4 100G-FR1	100G-LR4 100G-LR1	100G-ER4 100G-ER1



QSFP28 100G-LR1



QSFP28 100G-LR1



QSFP-DD 400G-PLR4



QSFP28 100G-LR1

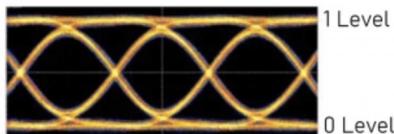
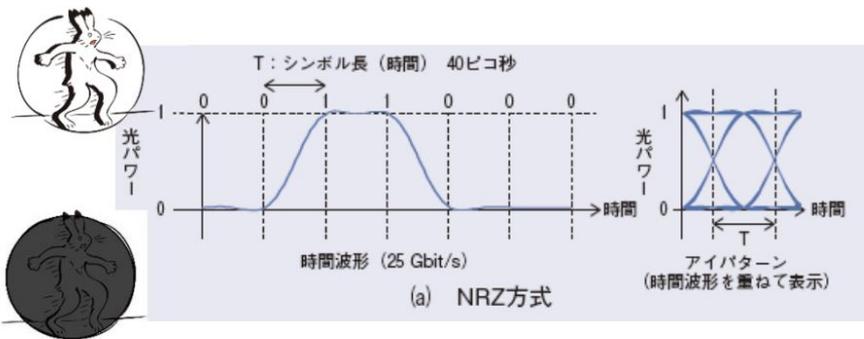
100G-LR4 / 100G-LR1 仕様比較

- 100G-LR1 では変調方式に PAM4 が採用されている
- PAM4 信号は SNR(信号雑音比)が低くノイズの影響を受けやすいため、100G-LR1では FEC(前方誤り訂正)が内蔵されている

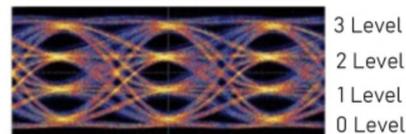
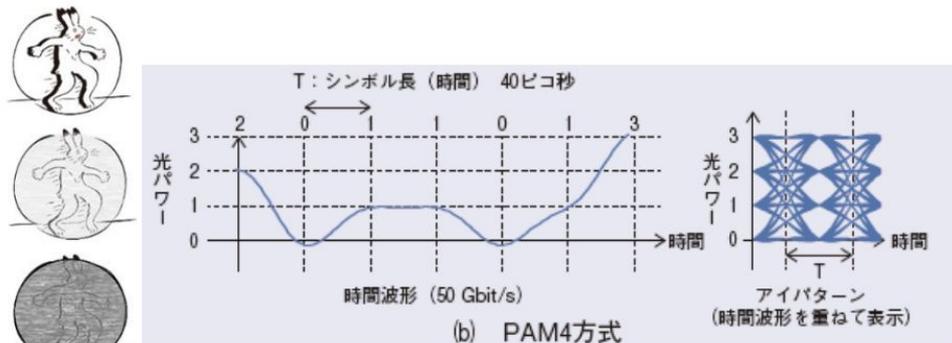
	100G-LR4	100G-LR1
伝送距離	10km	10km
ファイバ種別	SMF	SMF
波長数	25Gbit/s × 4波長	100Gbit/s × 1波長
変調方式	NRZ	PAM4
FEC	任意	必須

変調方式の違い

- NRZ (Non Return Zero)
 - 0,1 の2値で構成



- PAM4 (4-level Pulse Amplitude Modulation)
 - 0,1,2,3 の4値で構成



→ PAM4 方式では、NRZ 方式の2倍の情報量を伝送可能

株式会社ブロードバンドタワー, 光イーサネットの変化とオーバー100G対応,
<https://www.nic.ad.jp/ia/materials/iw/2018/proceedings/s05/s5-saeki.pdf>

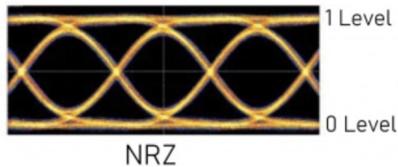
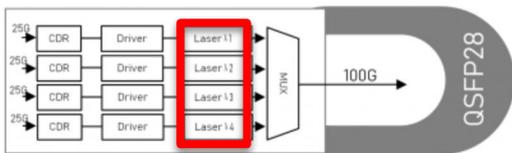
fiberroad, 100G光トランシーバーの規格,
<https://fiberroad.com/ia/resources/articles/edge-data-center-interconnection/>

100G-LR1 期待するメリット ①コスト削減

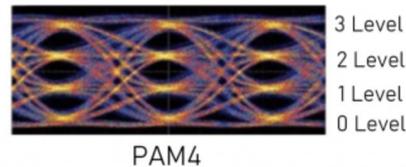
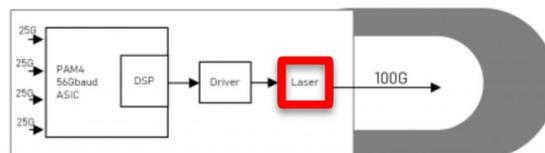
● 100G-LR4 より 100G-LR1 のほうが安価

- ○: 100G-LR1 は高価なレーザーが1個で済む
- △: 一方で、100G-LR1 では FEC や GEARBOX といったIC系が高価
- → **現状でも 100G-LR1 のほうが安価で今後も低価格化が進む可能性あり**
 - IC系は低価格化が早い傾向があるので、今後に期待
 - BigTech での導入が進めば一気に低価格化もあり得る ... ?

● 100G-LR4



● 100G-LR1



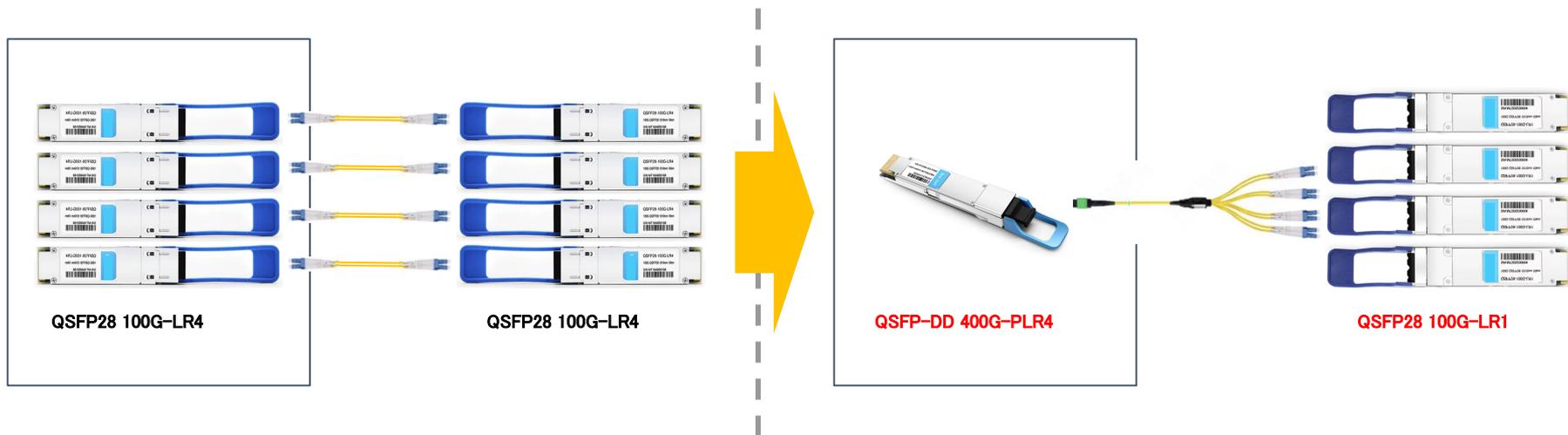
fiberroad, 100G光トランシーバーの規格.

<https://fiberroad.com/ja/resources/articles/edge-data-center-interconnection/>

100G-LR1 期待するメリット ②収容効率向上

- 400G 光モジュールを B/O して使うことで収容効率向上

- ○: ラックリソースの利用効率向上、集約側の急速なポート枯渇抑制、ケーブル逼迫緩和
- △: B/Oに必要な SMF MPO-LC ケーブルが高価、資産管理コスト増
- → **メリットのほうが圧倒的に大きく、リソースの活用には有効** と考える
 - 機器費用、ラックスペース、電力、増強稼働 etc.

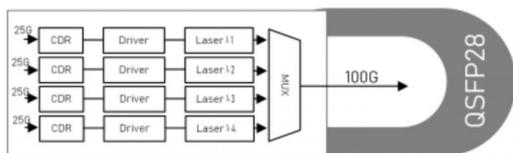


100G-LR1 期待するメリット ③故障リスク低下

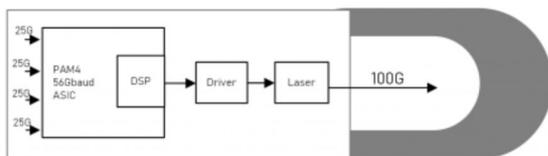
● MTBF (平均故障間隔) 低下

- 100G-LR4: 4つの波長を使用するため、対応する光学部品やレーザーが必要になり構造が複雑
- 100G-LR1: 単一波長で1つのレーザーを使用するため、コンポーネントが少なくシンプル
- → **100G-LR1 のほうが故障リスクが低くなることが期待されている**

● 100G-LR4

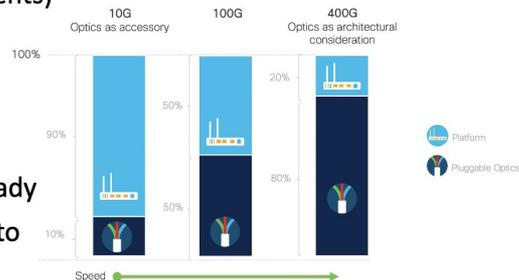


● 100G-LR1



100G LR1: Lower initial investment and TCO

- Reduce (optical) complexity (1 instead of 4 optical components)
- **Lower MTBF & MTBR**
- Backwards compatible and interoperable (switch/router, cabling, etc.)
- LR1 is about 20% cheaper, already
- High production volumes lead to further price erosion



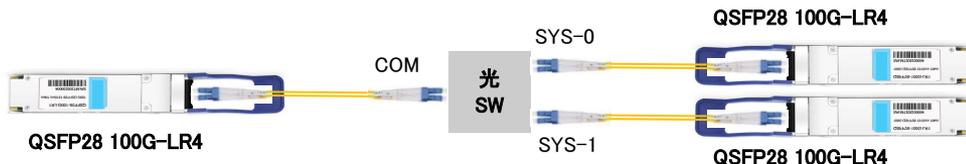
DE-CIX, Why 1 is sometimes more than 4 The merit of 100G LR1 vs LR4, https://peering-forum.eu/2023/wp-content/uploads/sites/6/2022/08/P3_Thomas-King_Why-1-is-sometimes-more-than-4.pdf

fiberroad, 100G光トランシーバーの規格, <https://fiberroad.com/ja/resources/articles/edge-data-center-interconnection/>

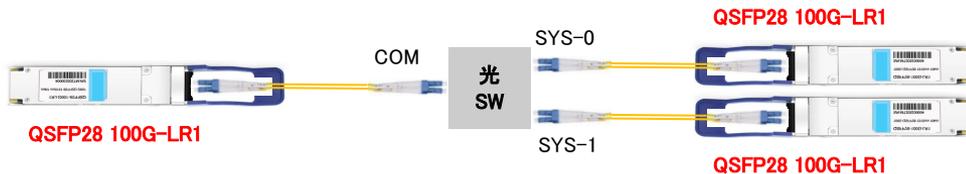
ちょっと検証やっていく

- 光スイッチ切り替えlink-up検証

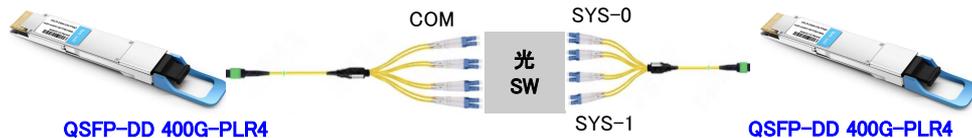
パターン① 100G-LR4同士



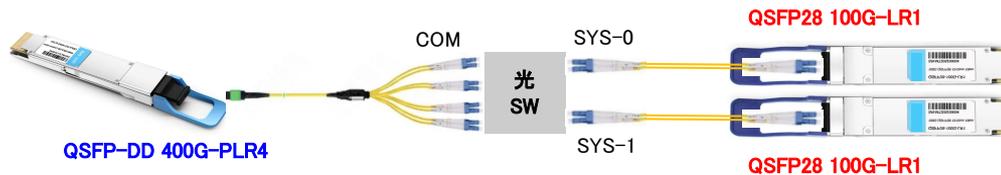
パターン② 100G-LR1同士



パターン③ 400G-PLR4同士



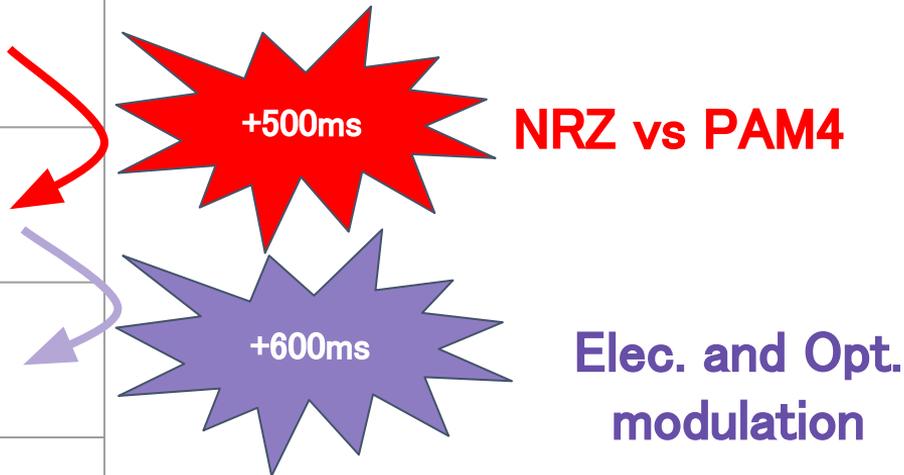
パターン④ 400G-PLR4 to 100G-LR1



検証による測定結果

- 各5回光スイッチ切り替え実施した際の切替時間

パターン	切替時間
① 100G-LR4同士	300ms
② 100G-LR1同士	800ms
③ 400G-PLR4同士	1400ms
④ 400G-PLR4 to 100G-LR1	1400ms

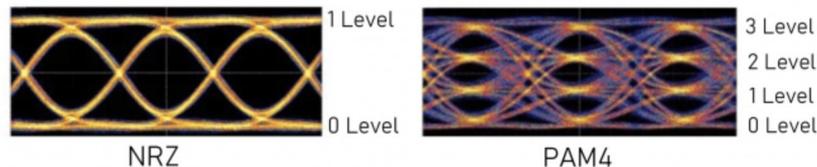
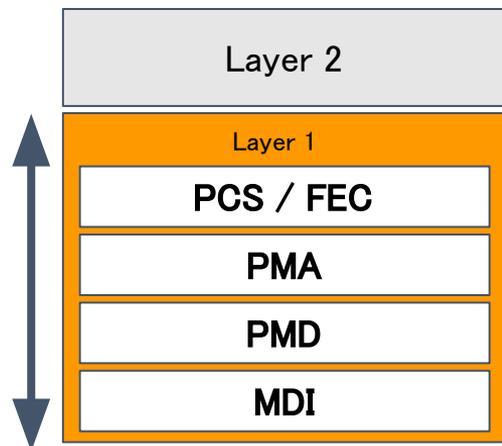


考察 (①100G-LR4 vs ②100G-LR1)

- 100G-LR1 の変調方式が PAM4 のため、NRZ より LinkUp に時間を要する

	100G-LR4	100G-LR1
変調方式	NRZ	PAM4
FEC	任意	必須

PHYのUPに
時間を要する



推測

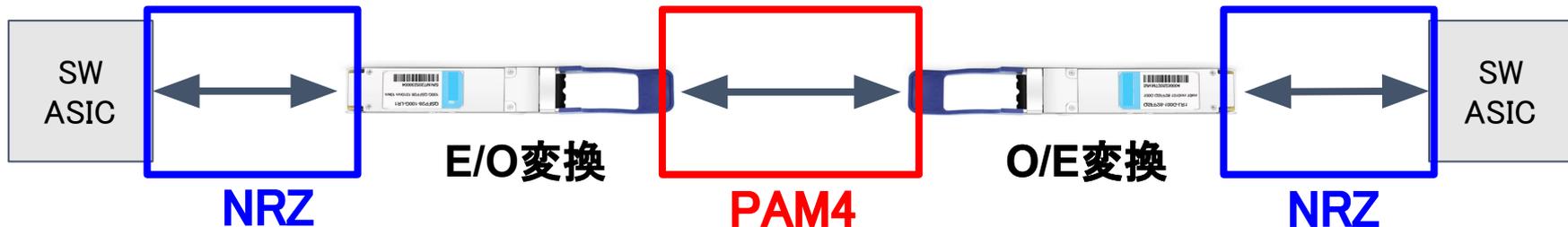
振幅が小さいのでエラーが乗り易い
FEC によるオーバーヘッドが NRZ と比較し大きい

IX相互接続実証実験を通じて見てきた400G導入で「変わること」「変わらないこと」, JANOG50 Meeting, <https://www.ianog.gr.jp/meeting/janog50/ix400/>
fiberroad, 100G光トランシーバーの規格, <https://fiberroad.com/ja/resources/articles/edge-data-center-interconnection/>
PAM4インターフェースの基礎(2), <https://www.macnica.co.jp/business/semiconductor/articles/broadcom/136684/>

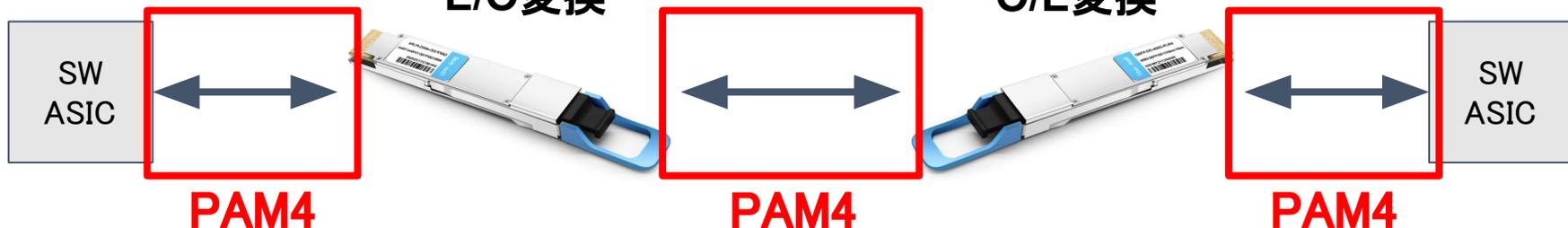
考察 (②100G-LR1 vs ③400G-PLR4)

- Optical interface 側が **PAM4** であってもElectical interface側が **NRZ** の可能性

100G-LR1



400G-PLR4



400Gイーサネットについて②, <https://arista.my.site.com/AristaCommunity/s/article/400gethernet2>

400Gbpsおよび800Gbpsへの道, <https://www.ianoz.gr.jp/meeting/ianoz46/wp-content/uploads/2020/06/ianoz46-shtsuchi-800g-00-%E5%9C%A7%E7%B8%AE%E6%B8%88%E3%81%BF.pdf>

100G-LR1 検討サマリ

● 期待するメリット

- ① コスト削減
 - 高価なレーザーが1個で済むため安価
 - 今後さらに低価格化していくことを期待(IC系の低価格化、需要増)
- ② 収容効率向上
 - 400G 光モジュールを B/O して使うことで集約側のリソースを有効活用
 - 機器費用、ラックスペース、電力、増強稼働 etc.
- ③ 故障リスク低下
 - 100G-LR4 に比べると構造がシンプルなため、故障リスク低下が期待されている

● 懸念点

- LinkUp に時間がかかる
 - 光スイッチあるから瞬断ですとは言い辛い値
 - 上位通信も考えるとさらに長期化

価格の低さ

100G-LR4

100G-LR1

収容効率の良さ

100G-LR4

100G-LR1

故障率の低さ

100G-LR4

100G-LR1

LinkUp時間の短さ

100G-LR4

100G-LR1

これからの100G接続

- 実際にあった某海外事業者からの声
 - 100G-LR4 高価 > 100G-LR1 安価
 - 100G-LR1 の方が Fail が起きづらい
 - **100G-LR4 での接続は対応できない、100G-LR1 でしか接続しない**

- 今後 400G 機器が増加していくにあたり、100G-LR1 での接続が増えるかも？
 - **100G 接続は、400G を B/O した 100G-LR1 しか用意できないと言われる日が来るかも**

- SFP112 の到来
 - 1波長 100G で QSFP より小さな SFP
 - 4波長では接続不可だったものが可能になってくる

サマリ

- **続く急速なトラフィック増加 → 収容逼迫の課題**
- **→ 課題解消に向けて 400G I/F を導入**
 - 収容効率向上、急速なポート枯渇抑制、ケーブルコスト削減 / 逼迫緩和

- **400G 時代の 100G 接続として、100G-LR1 を検討**
 - 400G 光モジュールを B/O して使うこともできる Single-lambda 製品
 - 100G-LR1 では変調方式に PAM4 を採用
 - 期待するメリット: ①コスト削減、②収容効率向上、③故障リスク低下
 - 懸念点: LinkUp 時間の増加

- **これからの 100G 接続**
 - 400G の普及が進むと共に、100G-LR1 での接続が増えるかも？
 - 既に 100G-LR4 での接続は対応できないという声もあり

“ The Internet ”

将来のために 100G-LR1 導入を推進しませんか？

- 400G I/F の導入計画はありますか？ / どのくらい進んでいますか？
- 100G-LR1 については検討していますか？
 - 100G-LR4 に比べて LinkUp 時間が増加すること
(= 光スイッチ切替時間増加)についてはどう考えますか？
 - 実際のところ、100G-LR1 での接続要望の声は？



No Peering, No Internet.