

IX相互接続実証実験を通じて見えてきた
400G導入で「変わること」「変わらないこと」

2022/7/14 (木) JANOG50@函館

鈴木 恒平 <kosuzuki@juniper.net>

JUNIPER
NETWORKS

Driven by
Experience™



Agenda

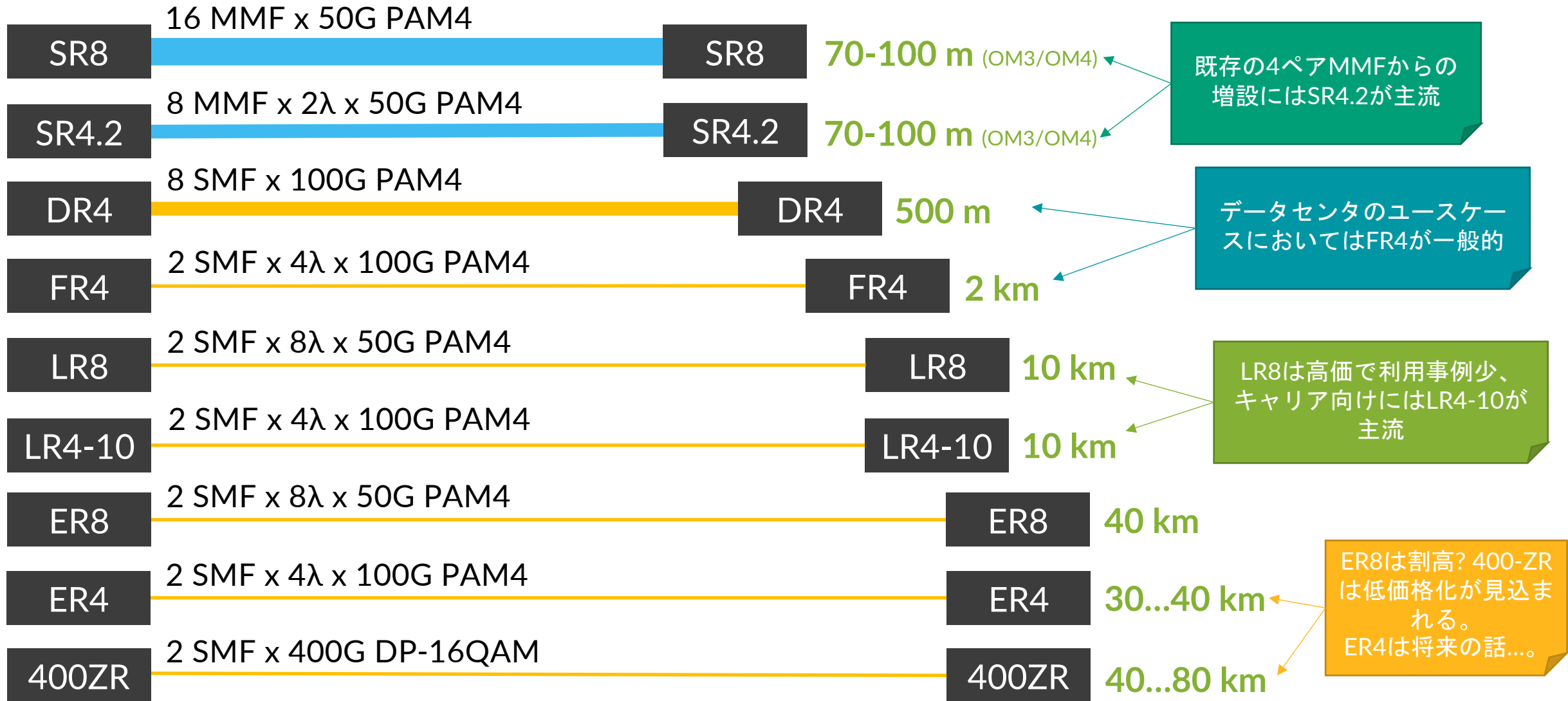
- 400G Optics のトレンドと装置側から見たトラブルシューティングの勘所
- Beyond 400G に向けて
- まとめ



Agenda

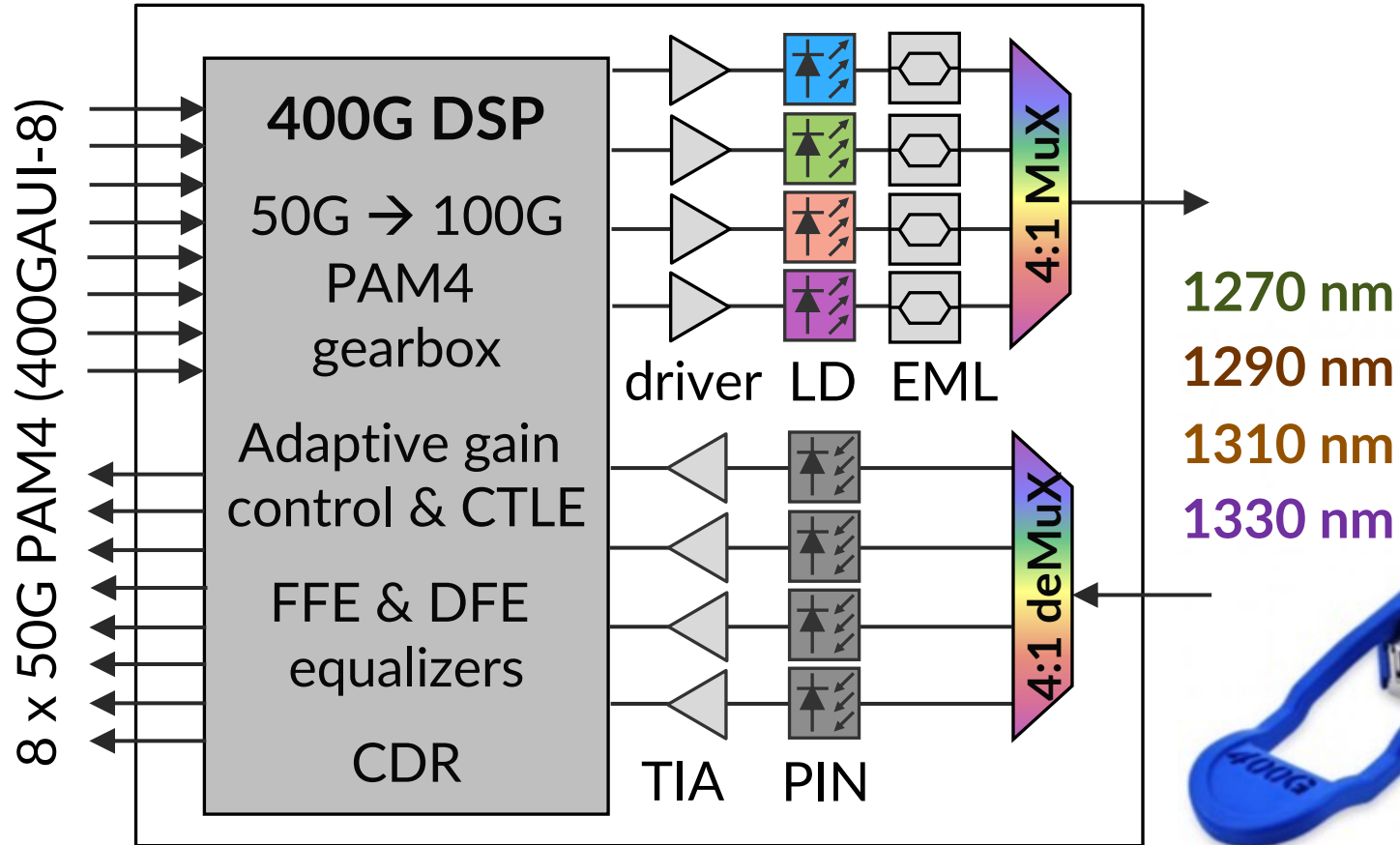
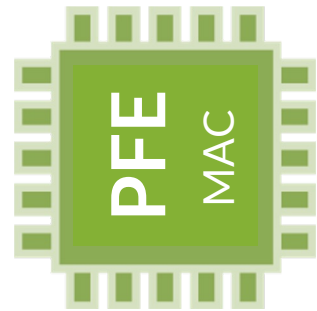
- 400G Optics のトレンドと装置側から見たトラブルシューティングの勘所
- Beyond 400G に向けて
- まとめ

400G Pluggable Optics の業界トレンド



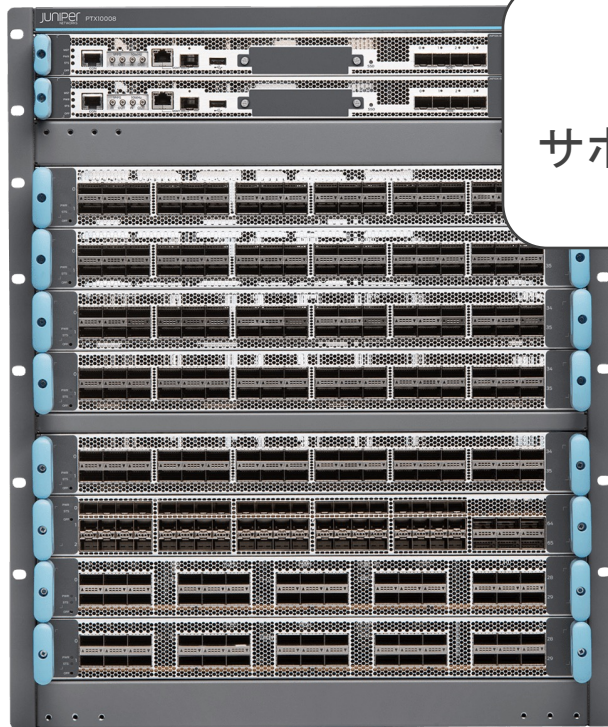
400G Pluggable Optics の内部構造

400G LR4 の場合



レーンあたり50G/100Gとなる次世代の Pluggable Optics は
デジタル信号処理を多用することで伝送損失を防いでいる

Opticsを挿してもInterfaceが上がらない...



Opticsを挿したけどリンクが上がらない。
サポート有無/OS versionは合っているはずなのに...



次の事柄についてチェックしてみよう

確認内容	チェック項目	被疑として考えられること
ルータ・スイッチが該当するインターフェースを認識しているか？	Yes/No	<ul style="list-style-type: none">機器の故障
配線が正しくされているか？	Yes/No	<ul style="list-style-type: none">ケーブルが断線しているミスケーブリング
対抗の機器間で同じOpticsを使っているか？	Yes/No	<ul style="list-style-type: none">Optics間の相性の問題
ケーブルのメディアタイプはOpticsに対して適切か？	SMF/MMF	<ul style="list-style-type: none">ケーブリングミス
ケーブル長は適切か？	Yes/No	<ul style="list-style-type: none">ケーブリングミス
ルータ・スイッチのコマンドラインでOpticsが認識されているか？	Yes/No	<ul style="list-style-type: none">Opticsの故障装置側のOSでその規格が未サポート
両端のTx/RxでOpticsのpowerが届いているか？	Yes/No	<ul style="list-style-type: none">Opticsの故障ケーブルの断線
show interface などの出力で表されるエラー種別は何か？	Local Fault / Remote Fault	<ul style="list-style-type: none">自装置における問題対向装置における問題
設定されたポートスピードは挿入したOpticsに対し適切か？	Yes/No	<ul style="list-style-type: none">設定ミス機器側の仕様 (明示的な速度設定が必要)
メーカー純正のOpticsでも同じような問題が発生するか？	Yes/No	<ul style="list-style-type: none">機器あるいはOpticsの被疑

機器側で確認できるOptics関連のステータス

show interfaces で確認可能な項目

- リンクステータス・LED
- PCS statistics (Bit errors / Errors blocks count)
- FEC mode, FEC statistics

その他

- SNMP Trap の有無
- Syslog message (Local Fault / Remote Fault など)

show interfaces diagnostics optics で確認可能な項目

- 波長ごとの Tx/Rx Power
- Optics の温度・消費電力
- バイアス電流
- 上記に関連したアラーム

これらの情報から統合的に判断してOpticsが動作しない理由を診断
アラームをトリガーするしきい値はOpticsの製造元によって設定される

400G相互接続再試験の結果 (2022/6/21-6/22)

Juniper MX10008 + Juniper 純正Optics の組み合わせでは現象は再現しなかったが、MX10008 + 3rd Party Optics の場合でのみリンクアップ遅延が発生 (リンクアップまでに約100秒かかる)

- リンクアップ遅延発生時には Juniper 側の機器で Local Fault / Remote Fault のログが繰り返し出現、後に安定し Link Up
- /var/log/messages にはリンクアップ遅延時に関連しそうなログは出現せず (90秒ほどログが出ない状態)
 - 原因は不明であるが、PHYが認識されるまでのハードウェア間での問題か？

QDD-400G-FR4	Arista 純正	Arista + 3rd Party Optics
Juniper 純正	現象再現せず	現象再現せず
Juniper + 3rd Party Optics	<u>リンクアップ遅延が再現</u>	試験未実施

3rd Party Optics の使用にあたって純正Opticsでは発生しない問題が見られた。
装置メーカーでの解析ができないリスクを許容できるかの判断が 3rd Party Optics 導入の勘所？



Agenda

- 400G Optics のトレンドと装置側から見たトラブルシューティングの勘所
- Beyond 400G に向けて
- まとめ

Beyond 400G に向けて...

IEEE での 800GE & 1.6TE Ethernet の標準化動向



40GE & 100GE
10G & 25G electrical I/O

IEEE 802.3ba-2010

Standard start: July 2006
Standard approved: June 2010

200GE & 400GE
50G & 100G electrical I/O

IEEE 802.3bs-2017

Standard start: March 2013
Standard approved: Dec. 2017

800GE & 1.6TE
100G & 200G electrical I/O

IEEE P802.3df

200 Gb/s, 400 Gb/s, 800 Gb/s, and
1.6 Tb/s Ethernet Task Force

Standard start: October 2020
Standard approved: Approx. 2025

新たなEthernet標準には新しい技術が必要となる
技術的可能性の検討・製品レベルでの実現に向けて長期的なコンセンサスの形成がキーポイント

800G Ethernet の市場投入

ETC specification for 800GE

DC事業者を始めとして800Gポートの導入が期待されている

- Broadcom Tomahawk 4 などのスイッチシリコンを使用
- 当初は主に2 x 400GEとして使用されることが想定

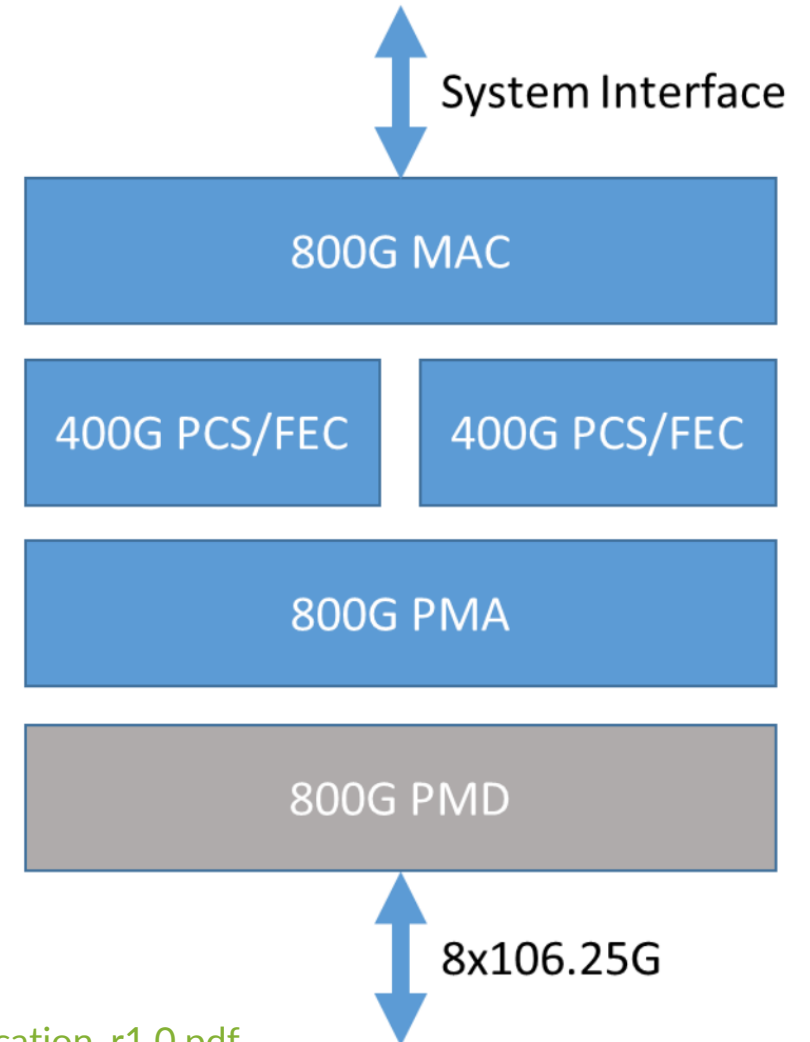
Ethernet Technology Consortium (ETC)* が800ギガビットイーサネットの仕様を公開 (2020年4月)

- ETC仕様では、400GBASE-RのPCS/FEC仕様を再利用することで800Gに対応

800GEポートに対応したルータやスイッチは2023-24年には市場に流通することが予測されている

* 800GbE仕様発表と同時に 25 Gigabit Ethernet Consortium から名称変更

https://ethernettechnologyconsortium.org/wp-content/uploads/2020/03/800G-Specification_r1.0.pdf



800G Optics のフォームファクタ

QSFP-DD800

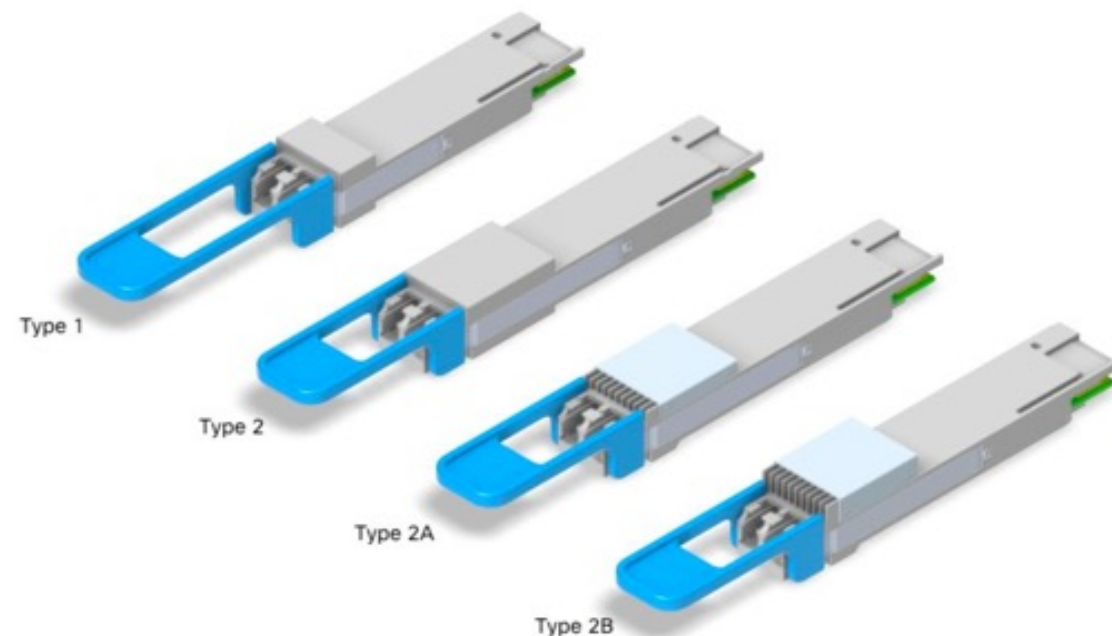
QSFP-DD800

QSFP-DD800 MSA は 800G Optics に対応するために QSFP-DD 仕様を拡張

- QSFP28 (100G) および QSFP-DD (400G) との下位互換性を維持

800G optics の消費電力

- DR8/2xFR4クラスでは 14~16W (400G Optics より若干増加)
- 800G-ZR/ZR+ではポートあたり25W
- QSFP-DD を type 2B モジュールにすることで25Wへの対応を実現



<http://www.qsfp-dd800.com/wp-content/uploads/2020/03/QSFP-DD-800-Hardware-1p0-3-6-20%20FINAL.pdf>

QSFP-DD800ではポートあたり最大25Wまで対応
同一フォームファクタで800Gポートに加えて、コヒーレントDWDM光モジュール(ZR/ZR+)に対応



Agenda

- 400G Optics のトレンドと装置側から見たトラブルシューティングの勘所
- Beyond 400G に向けて
- まとめ

まとめ

400Gでも通常のOpticsのデバッグと変わらない (1G/10G/40G/100G)

- 物理層から疑うことは400Gでも変わらない
- 3rd Party Optics を利用している場合は、メーカー純正のOpticsで同じ問題が起こるかを確認

400Gに限らず、伝送規格についての理解は必須

- 規格と距離、使用すべきファイバーの種別(メディアタイプ)の理解はMUST
- 100Gと同様にレーンごとに見る必要がある
- 400Gは規格が多いため、使う規格とそれがどのような波長を使っているかは特に注意

800G/1.6Tは標準化段階だが、早い段階で仕様を策定している団体もある

- Ethernet Technology Consortium による 800 Gigabit Ethernet (GbE) Specification
- IEEEによる800G/1.6Tは標準化の最中 (標準化完了は2025年ごろ?)

400GEは商用利用可能な水準。オペレーターによる運用知見の蓄積はこれから。



Thank you

JUNIPER
NETWORKS | Driven by
Experience™