

リンクアップするってどういうこと？ について考えるBoF

Janog56 BoF

2025/7/30

マクニカクラブスカンパニー

阿部野一郎

発表者紹介



阿部野 一郎

株式会社マクニカ
クラビスカンパニー

略歴：

2003年よりマクニカにてワイヤレス系製品のFAEとして活動
2020年くらいから光通信の分野にも活動の範囲を拡げています

質問は途中いつでも挟んでもらってOKです！

リンクアップ・・・

リンクアップする・しない

リンクアップするけど、フラップする・しない

そもそもリンクアップする・しないって・・・

誰が何を見て判断してるの？！

誰が

- ・さすがに光トランシーバじゃなさそう
- ・装置側なんだろうけど・・・

何を？

- ・光の受信のレベル？
- ・Bit Error Rate？
- ・受信データの中身？

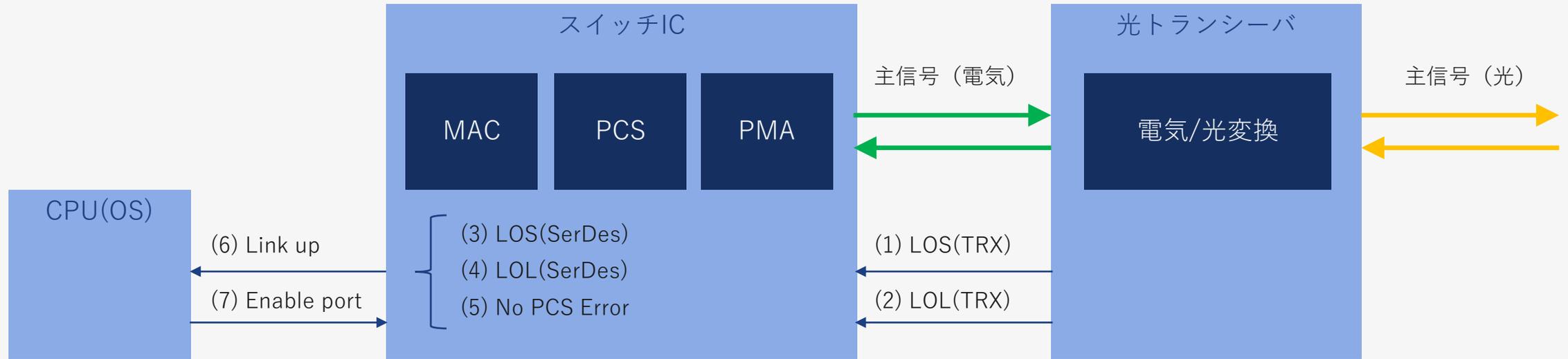
リンクアップ・・・・

そもそもリンクアップする・しないって・・・・

誰が何を見て判断してるの？！

調べてみて分かったことを、共有します

リンクアップするまでの大まかな流れ(1)



(6) SerDesのLink upが立っている
(7) 該当ポートがEnable

(6)&(7)=True → リンクアップ
* 装置のリンクアップLEDが点灯

SerDesの
(3) Loss Of Signalが無い
(4) Loss Of Lockが無い

(5) PCSでのエラーが無い

(3)&(4)&(5)=True → (6) Link upを上位に通知

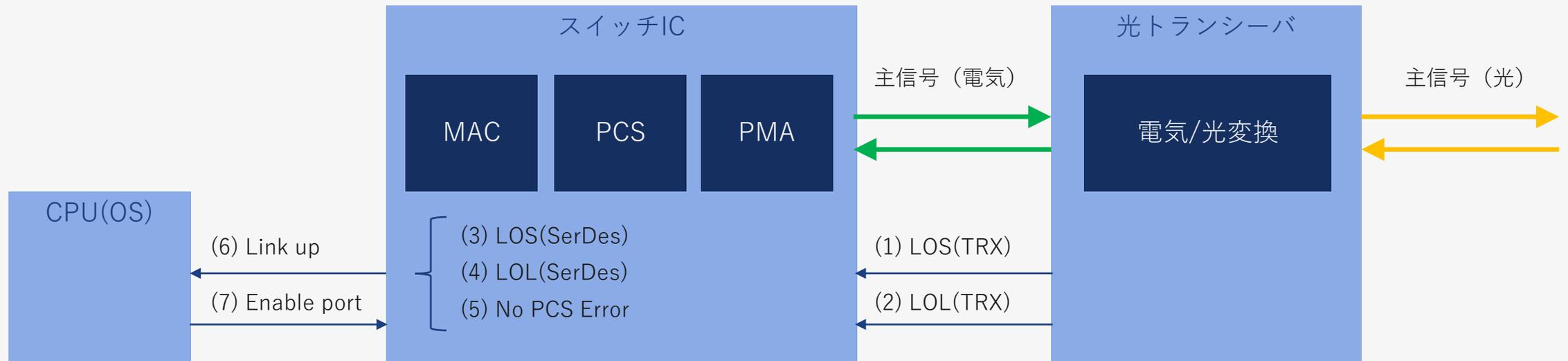
* (1)と(2)も条件である可能性あり

光トランシーバの
(1) Loss Of Signalが無い
(2) Loss Of Lockが無い

(1)と(2)はレジスタにてI2C経由でスイッチICに通知される

* 通知先はCPUかも？

リンクアップするまでの大まかな流れ(2)



(6) SerDesのLink upが立っている
(7) 該当ポートがEnable

(6)&(7)=True → リンクアップ
* 装置のリンクアップLEDが点灯

SerDesの
(3) Loss Of Signalが無い
(4) Loss Of Lockが無い
(5) PCSでのエラーが無い

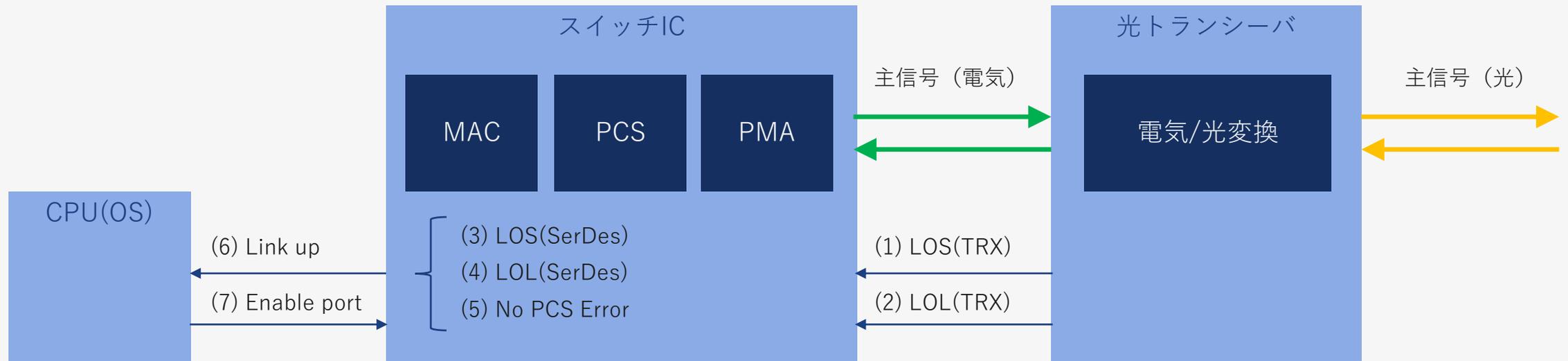
(3)&(4)&(5)=True → (6) Link upを上位に通知
* (1)と(2)も条件である可能性あり

光トランシーバの
(1) Loss Of Signalが無い
(2) Loss Of Lockが無い

(1)と(2)はレジスタにてI2C経由でスイッチICに通知される

* 通知先はCPUかも？

リンクアップするまでの大まかな流れ(3)



(6) SerDesのLink upが立っている
(7) 該当ポートがEnable

(6)&(7)=True → リンクアップ
* 装置のリンクアップLEDが点灯

SerDesの
(3) Loss Of Signalが無い
(4) Loss Of Lockが無い

(5) PCSでのエラーが無い

(3)&(4)&(5)=True → (6) Link upを上位に通知

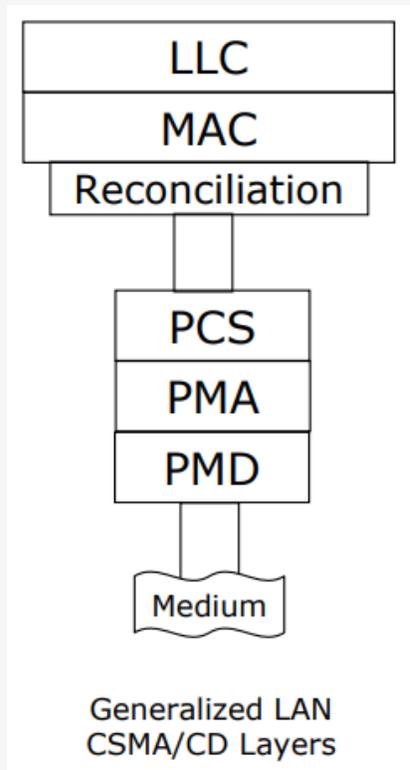
* (1)と(2)も条件である可能性あり

光トランシーバの
(1) Loss Of Signalが無い
(2) Loss Of Lockが無い

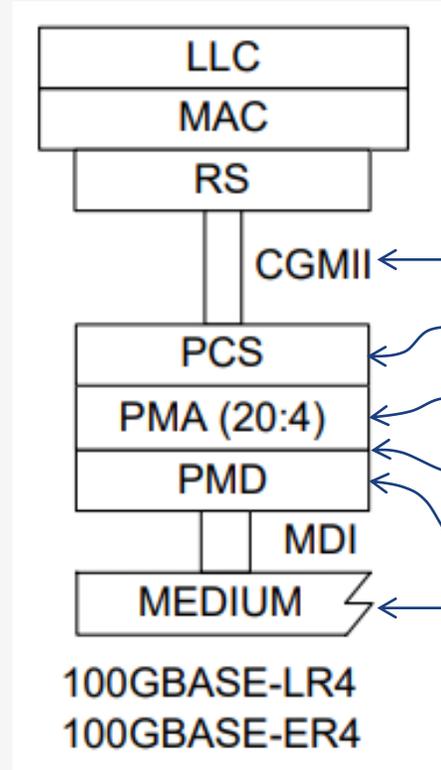
(1)と(2)はレジスタにてI2C経由でスイッチICに通知される

* 通知先はCPUかも？

PCSで何をやってるか (1)



一般化したレイヤ構成



(100G) 25G x 4レーン構成

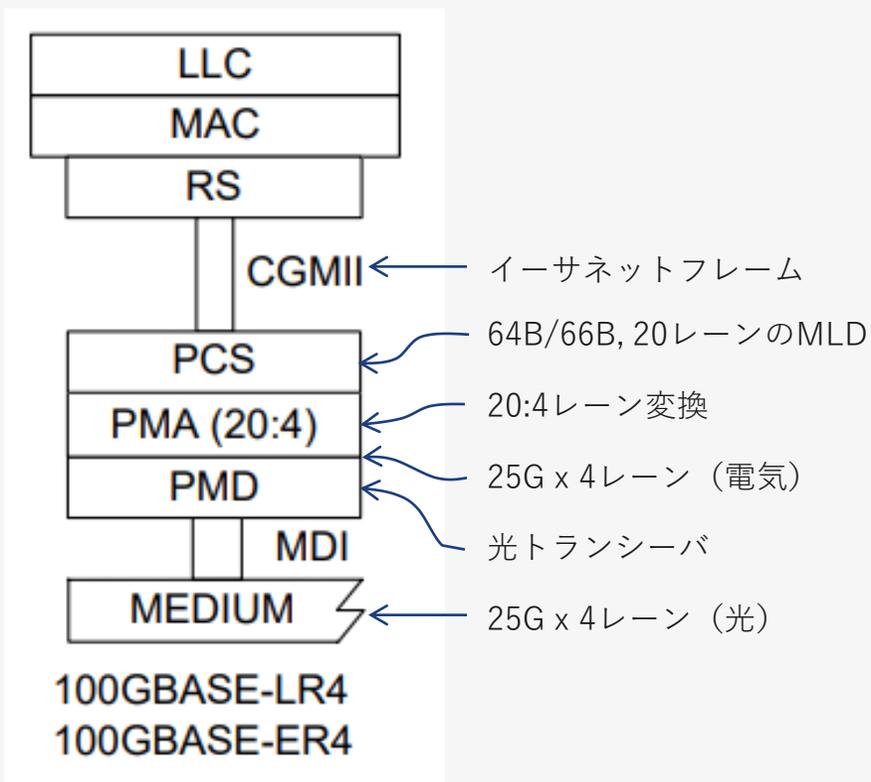
CGMII ← イーサネットフレーム
 PCS ← 64B/66B, 20レーンのMLD
 PMA (20:4) ← 20:4レーン変換
 PMD ← 25G x 4レーン (電気)
 MDI ← 光トランシーバ
 MEDIUM ← 25G x 4レーン (光)

- LLC Logical Link Control
- MAC Media Access Control
- RS Reconciliation Sub-layer
- PCS Physical Coding Sub-layer**
- PMA Physical Media Attachment
- PMD Physical Media Dependent
- MDI Medium Dependent Interface

- MLD Multi Layer Distribution

出典 : IEEE P802.3ba Architecture Overview
https://www.ieee802.org/3/time_adhoc/public/apr09/dambrosia_03_0509.pdf

PCSで何をやってるか (2)



(100G) 25G x 4レーン構成

出典：IEEE P802.3ba Architecture Overview
https://www.ieee802.org/3/time_adhoc/public/apr09/d_ambrosia_03_0509.pdf

PCS

送信：イーサネットフレームを64ビットごとに分割し、2ビットを足して、66ビットのブロックを作る
これらを仮想的に20本のレーンに分けて送る (MLD)

受信はこの逆

(5) PCSエラーが無い・・・ちゃんとイーサネットフレームに戻せている

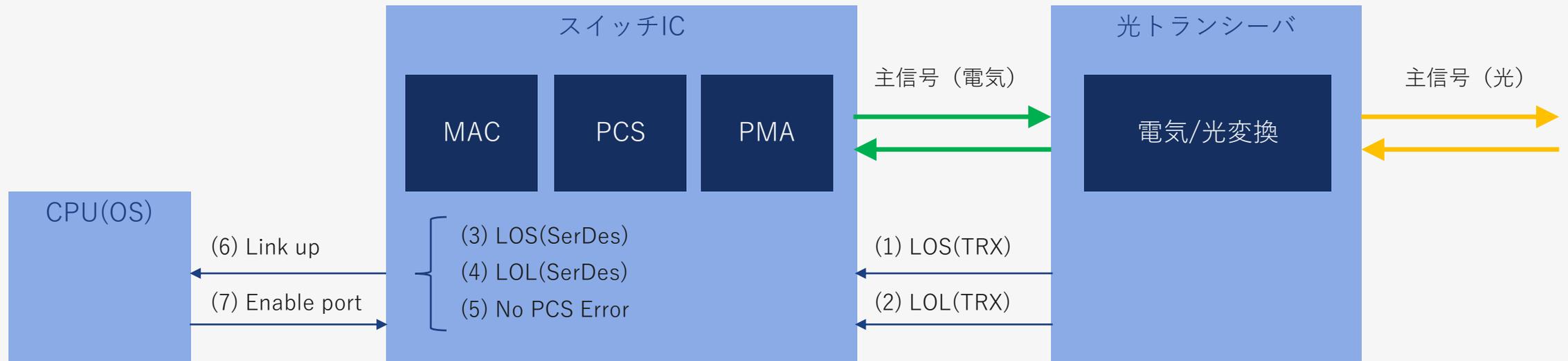
PMA

送信：20レーンを4レーンにMuxし、25G x 4レーンとする
これを電気→光変換するのが光トランシーバ

受信はこの逆

参考：100ギガビットイーサネットについて (IJJ)
https://www.ijj.ad.jp/dev/report/iir/pdf/iir_vol13_network.pdf

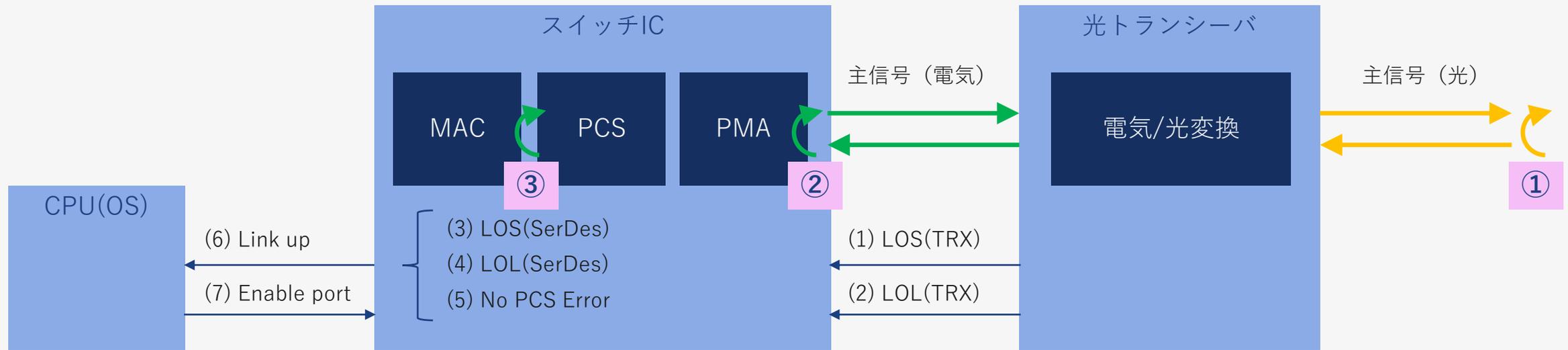
リンクアップがうまくいかないときの切り分け（フラグやエラーチェック）



(1)~(7)のうち、どれに問題があるか？
(1)と(2)は光トランシーバから
(3)~(5)はスイッチICから

CLIでチェックできる？！

リンクアップがうまくいかないときの切り分け（ループバック）



③ Deep Loopback
Shallow Loopbackとの組み合わせで、PMA/PCSに問題があるかどうかの切り分けができる

② Shallow Loopback
受信信号（パケット）をそのまま返すこれでリンクアップするかどうか？
する：受信パケットに問題ない
しない：受信パケットに問題あり

① Optical Loopback
光のまま返してみる

各種ループバックの組み合わせ → どこに問題があるかの切り分けに役立つ
Shallow/Deep Loopbackができるかどうかは、装置やOSによるのでは？！

おまけ

物理層でのチェックポイント

電気信号はキレイ？（装置と光トランシーバ間）
光信号はキレイ？（光トランシーバ同士）

ポート依存性？
光トランシーバや光ファイバを交換したらどうなる？
光ファイバの種類は正しい？

物理的な接続点の汚れ、破損、勘合具合

その他には・・・

ベンダーロックとか

MACNICA

- ・本資料に記載されている会社名、商品またはサービス名等は各社の商標または登録商標です。なお、本資料中では、「™」、「®」は明記していません。
- ・本資料のすべての著作権は、第三者または株式会社マクニカに属しており、(著作権法で許諾される範囲を超えて) 無断で本資料の全部または一部を複製・転載等することを禁じます。
- ・本資料は作成日現在における情報を元に作成されておりますが、その正確性、完全性を保証するものではありません。

© Macnica, Inc.