

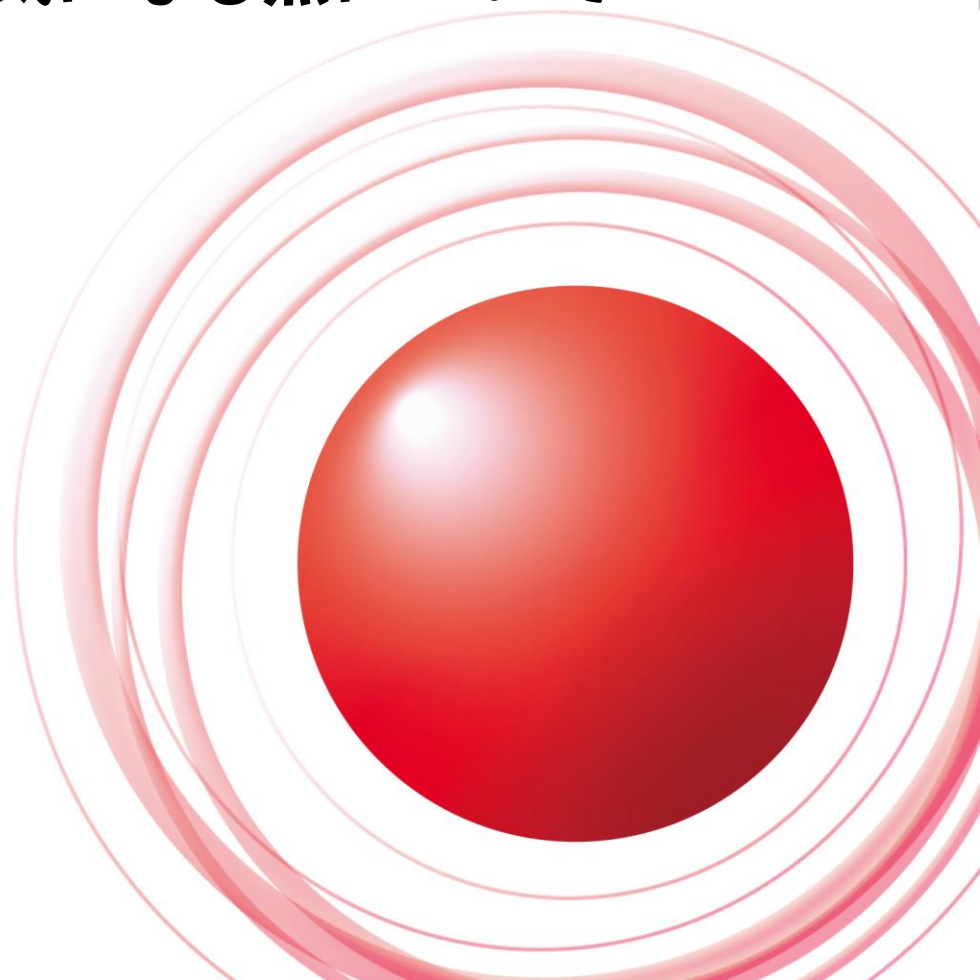
IXと100Gbit Ethernet

-運用を想定した場合に気になる点について-

IIJ Internet Initiative Japan

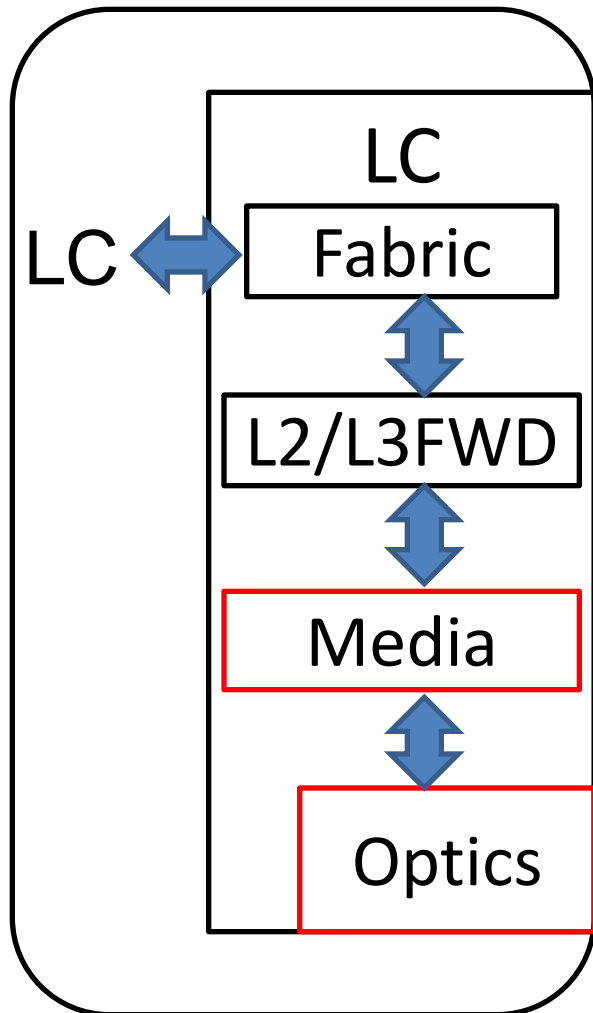
株式会社インターネットイニシアティブ
大内宗徳

Ongoing Innovation



100GEの特徴に基づく障害切り分け

ルータ/スイッチ



障害ポイント

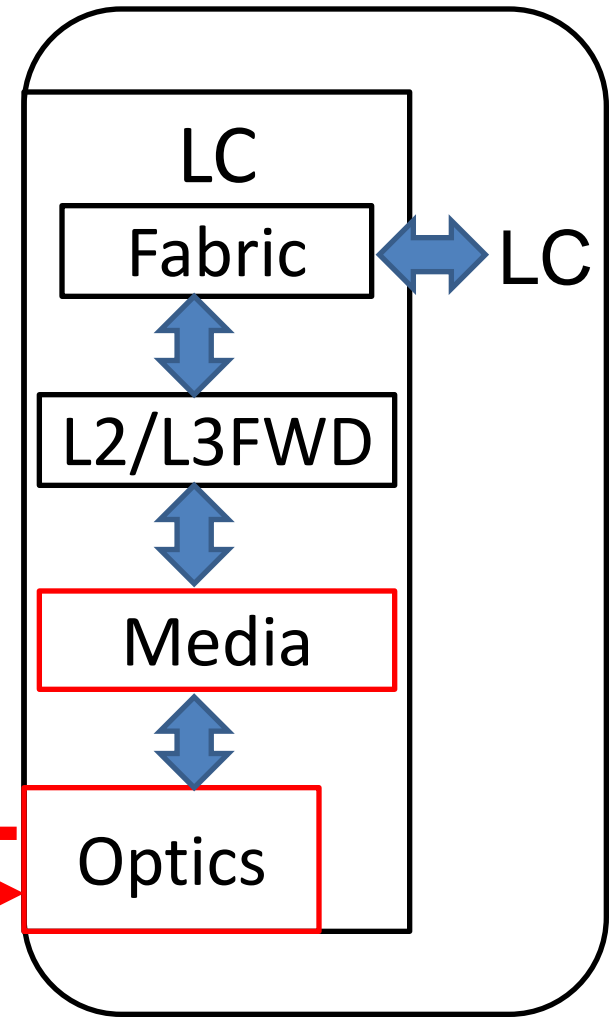
- 回線
- Optics
- LCのASIC
- LCのメモリ
- その他

MLD

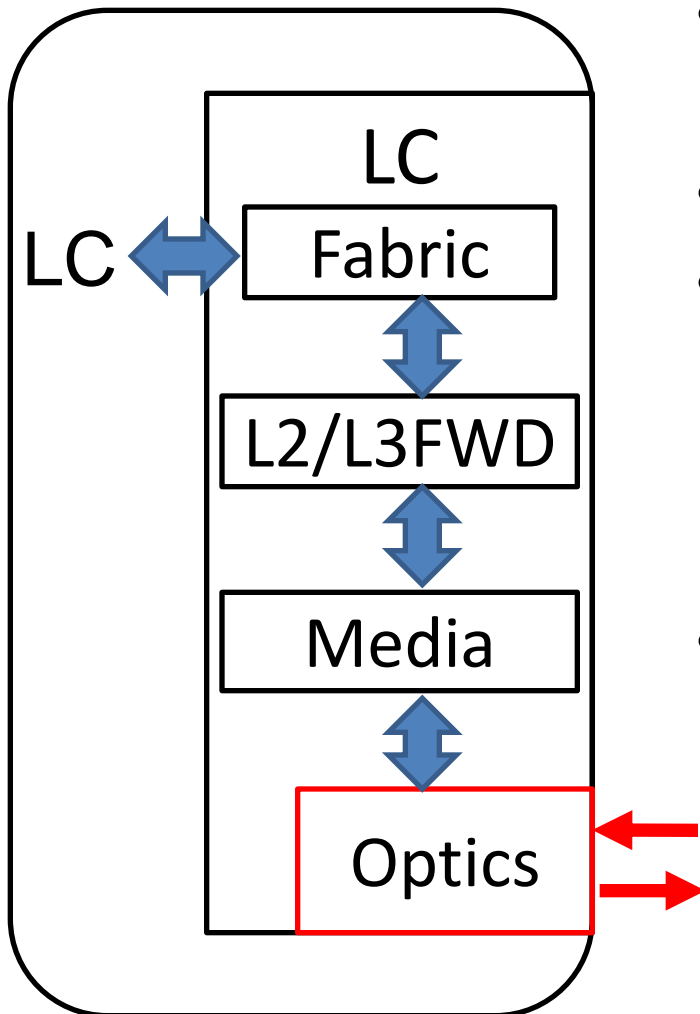
(Multi Lane Distribution)

WDM/ファイバ多重

ルータ/スイッチ



回線/Optics障害 送受光レベルでの切り分け



- 切り分けのために
 - 受光レベル測定が必要
- 一方100GBASE-LR4は4波長WDM
- 普通のパワーメータを使うと
 - 4波合計の受光レベルのみ測定
 - 波長毎に受光レベルの確認ができない！
- 対応方法
 - 波長毎測定可能なパワーメータ
 - CFPの波長単位の受光レベル測定機能

CFPの波長単位の送受光レベル測定機能

- **実装状況(試験時点で)**
 - Brocade → 開発中
 - Cisco → 合成/波長単位の両方
 - Juniper → 波長単位の受光レベルのみ

CFPの受光レベル表示機能について例 (Cisco)

```
RP/0/RP0/CPU0:cisco_CRS#sh controllers hundredGigE 0/0/0/0 phy  
CFP Detail Information:
```

Network lane A/D value:

```
Lane 0 Tx power: 1.0076 mW ( 0.0 dBm)  
Lane 1 Tx power: 1.1694 mW ( 0.7 dBm)  
Lane 2 Tx power: 0.9728 mW (-0.1 dBm)  
Lane 3 Tx power: 1.0842 mW ( 0.4 dBm)
```

```
Lane 0 Rx power: 1.1181 mW ( 0.5 dBm)  
Lane 1 Rx power: 1.2432 mW ( 0.9 dBm)  
Lane 2 Rx power: 1.2216 mW ( 0.9 dBm)  
Lane 3 Rx power: 1.1343 mW ( 0.5 dBm)
```

```
Total Tx power : 4.2340 mW ( 6.3 dBm)  
Total Rx power  : 4.7172 mW ( 6.7 dBm)
```

(参考)

Total powerから
波長単位のpowerを
概算したい場合

(Total mW) / 4

=>波長単位のmW

(Total dBm) - 6

=>波長単位のdBm

CFPの受光レベル表示機能について例 (Juniper)

```
lab@isp-t1600> show interfaces diagnostics optics
```

```
Physical interface: et-1/0/0:0
```

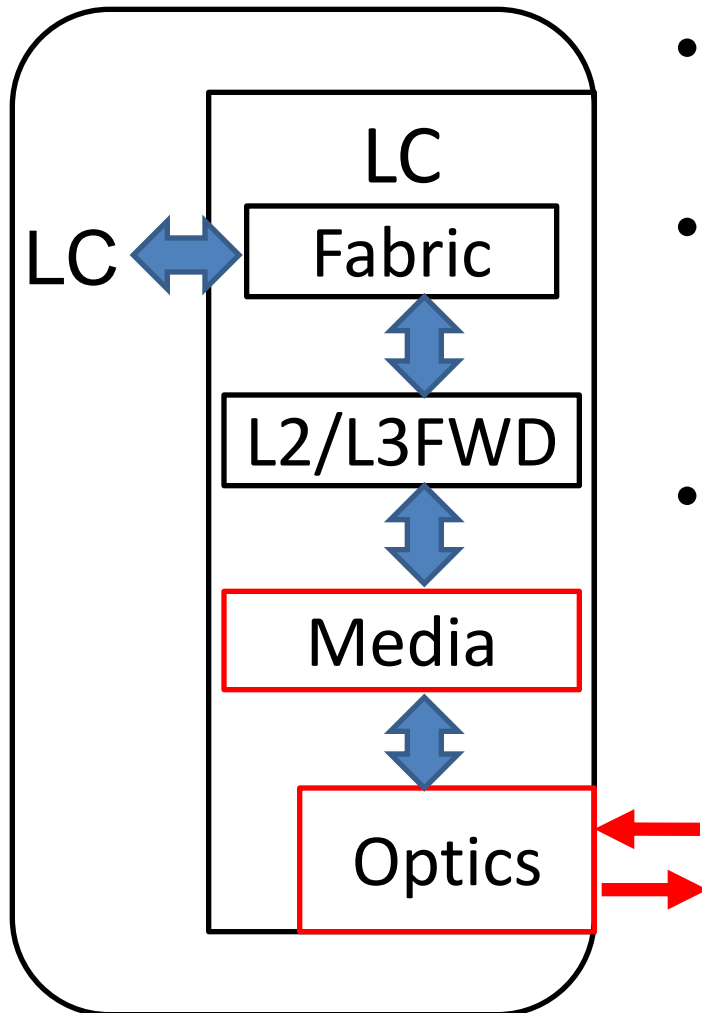
```
·  
Lane 0
```

```
Laser bias current           : 42.198 mA  
Laser output power          : 0.943 mW / -0.25 dBm  
Laser temperature           : 48 degrees C / 119 degrees F  
Laser receiver power        : 0.552 mW / -2.58 dBm
```

```
·  
Lane 1
```

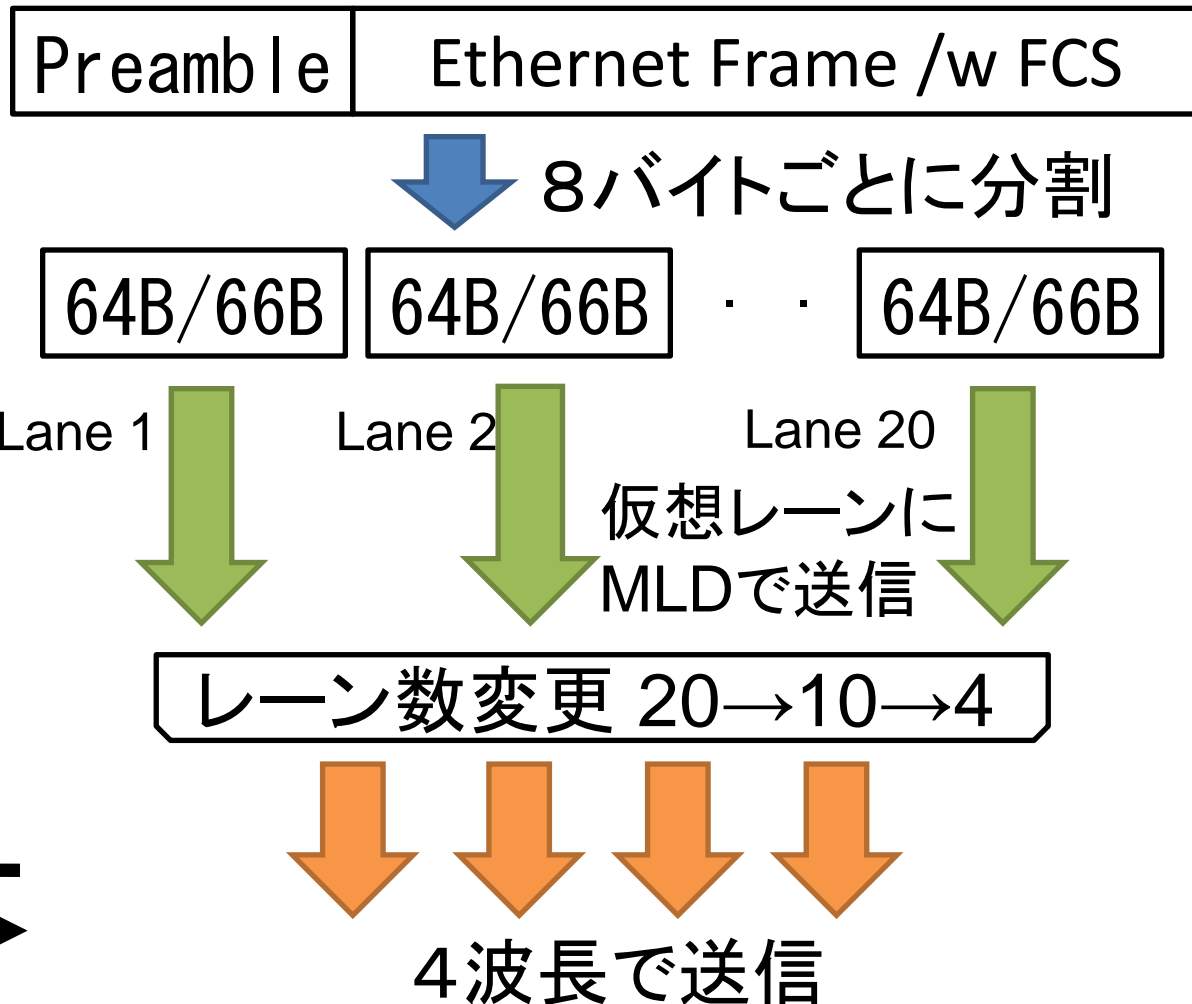
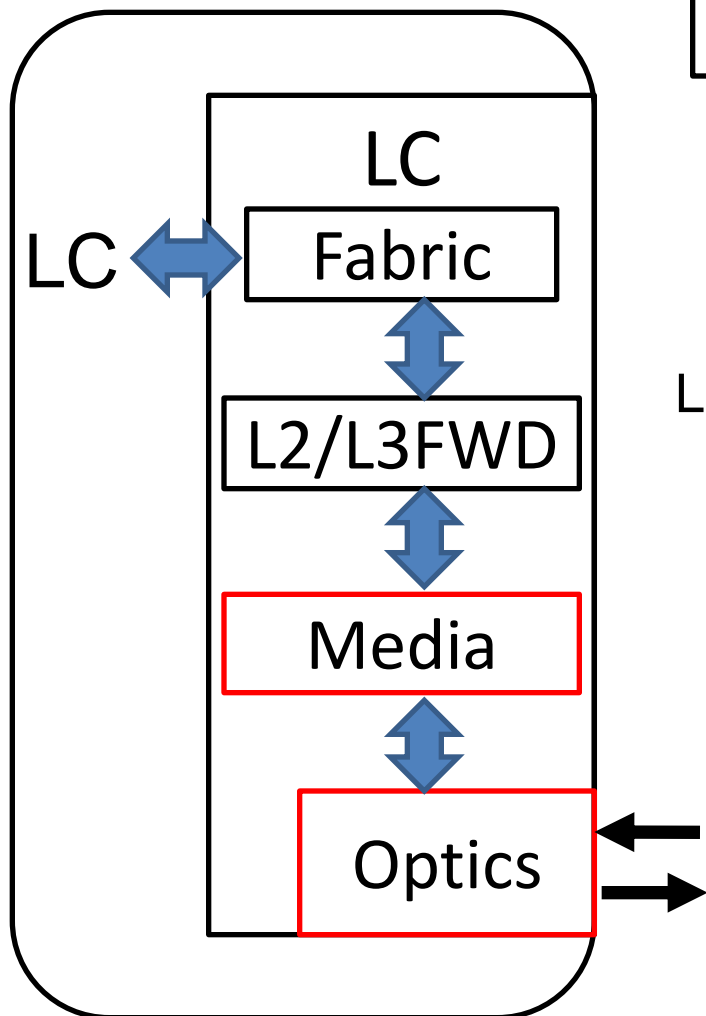
```
Laser bias current           : 41.832 mA  
Laser output power          : 0.980 mW / -0.09 dBm  
Laser temperature           : 50 degrees C / 123 degrees F  
Laser receiver power        : 0.567 mW / -2.47 dBm
```

回線/Optics/LCのASIC障害 パケットロスがある場合

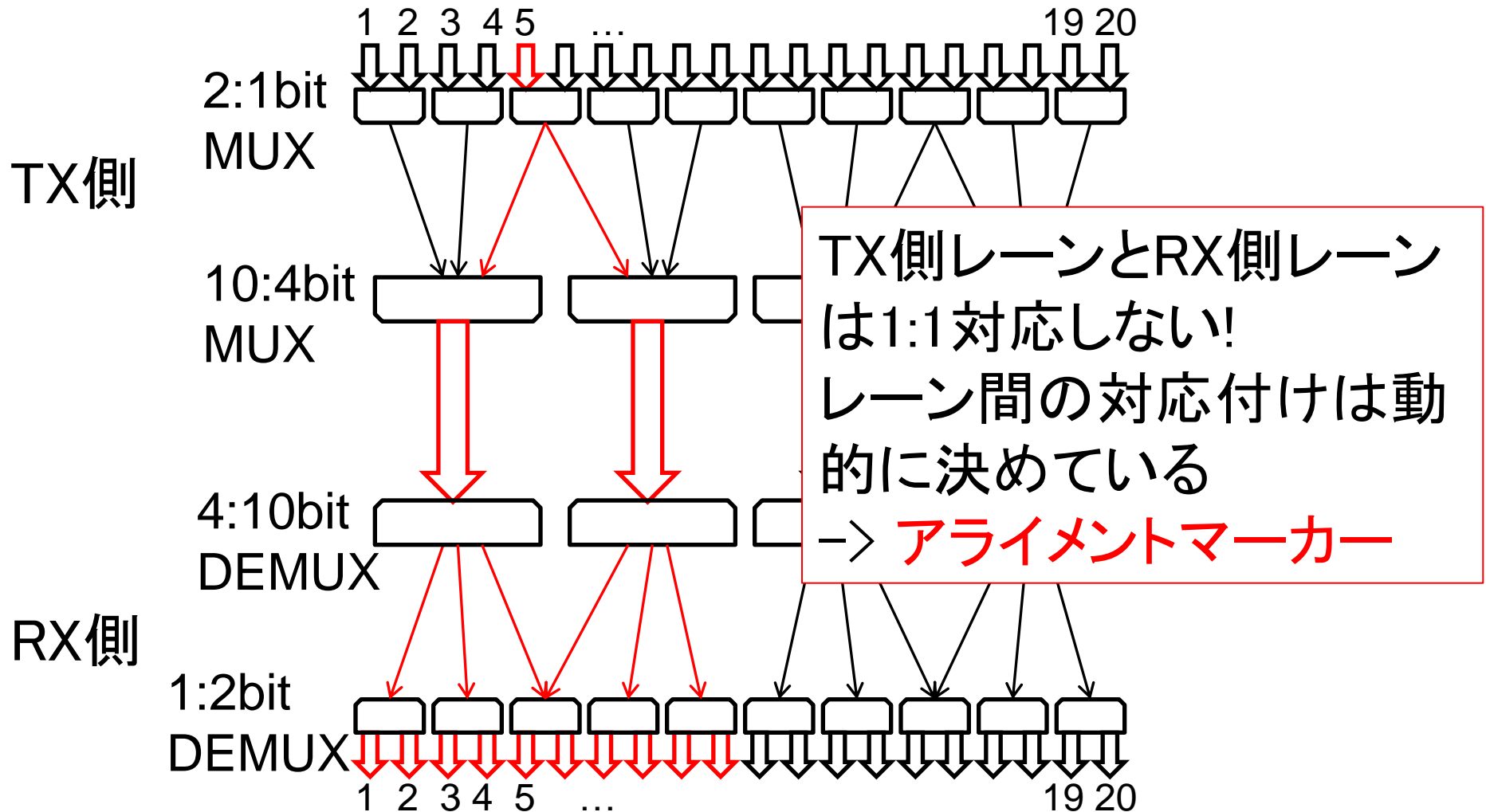


- 切り分けのためには
 - EthernetのCRCエラーの有無
- 100GEでは複数経路で伝送
 - 各経路のエラー監視ができないと障害切り分けが困難…
- 対応方法
 - 100GEでは仮想レーン単位でビットエラー発生を監視する機能が
 - BIP(Bit Interleaved Parity)
 - 詳細は次のスライドで…

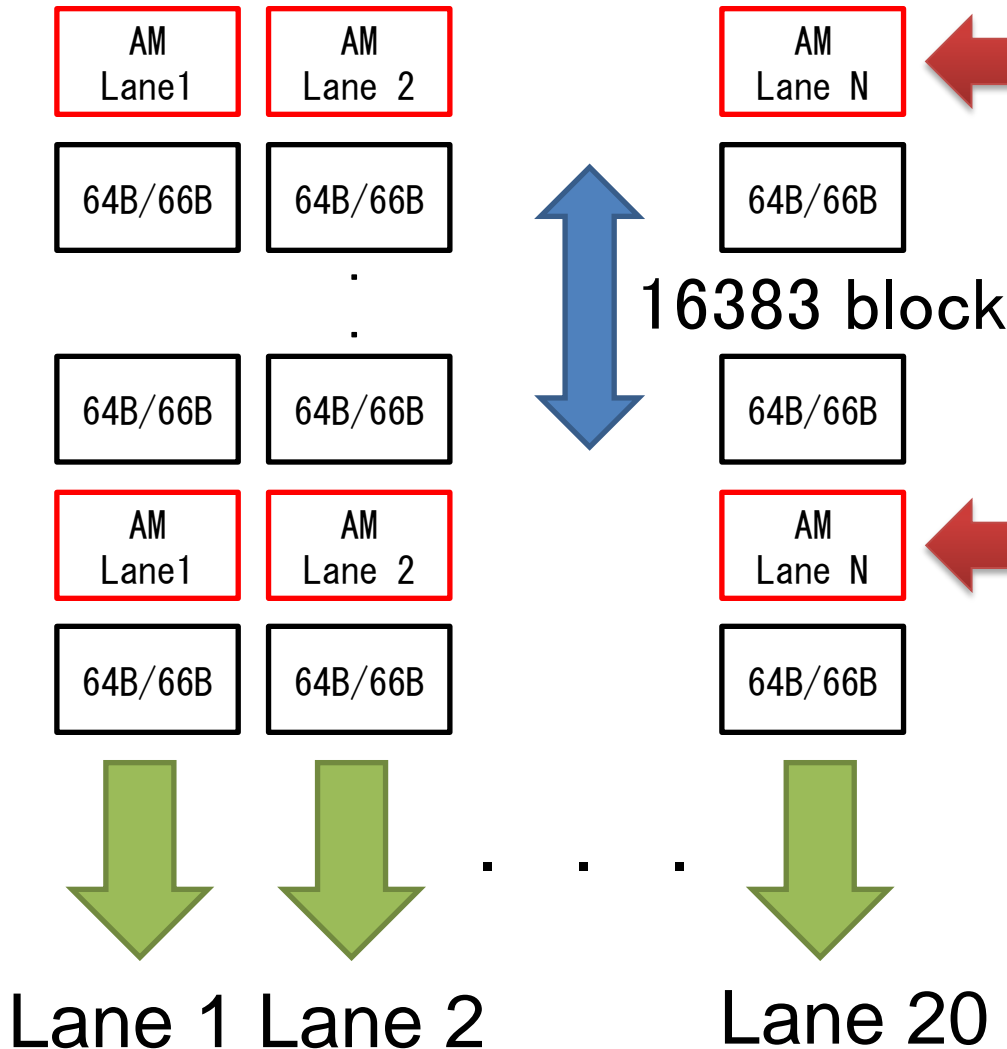
100GBASE-LR4のデータ送信の ざっくりとした仕組み



仮想レーン例 (100GBASE-LR4)



アライメントマーカ―(AM)

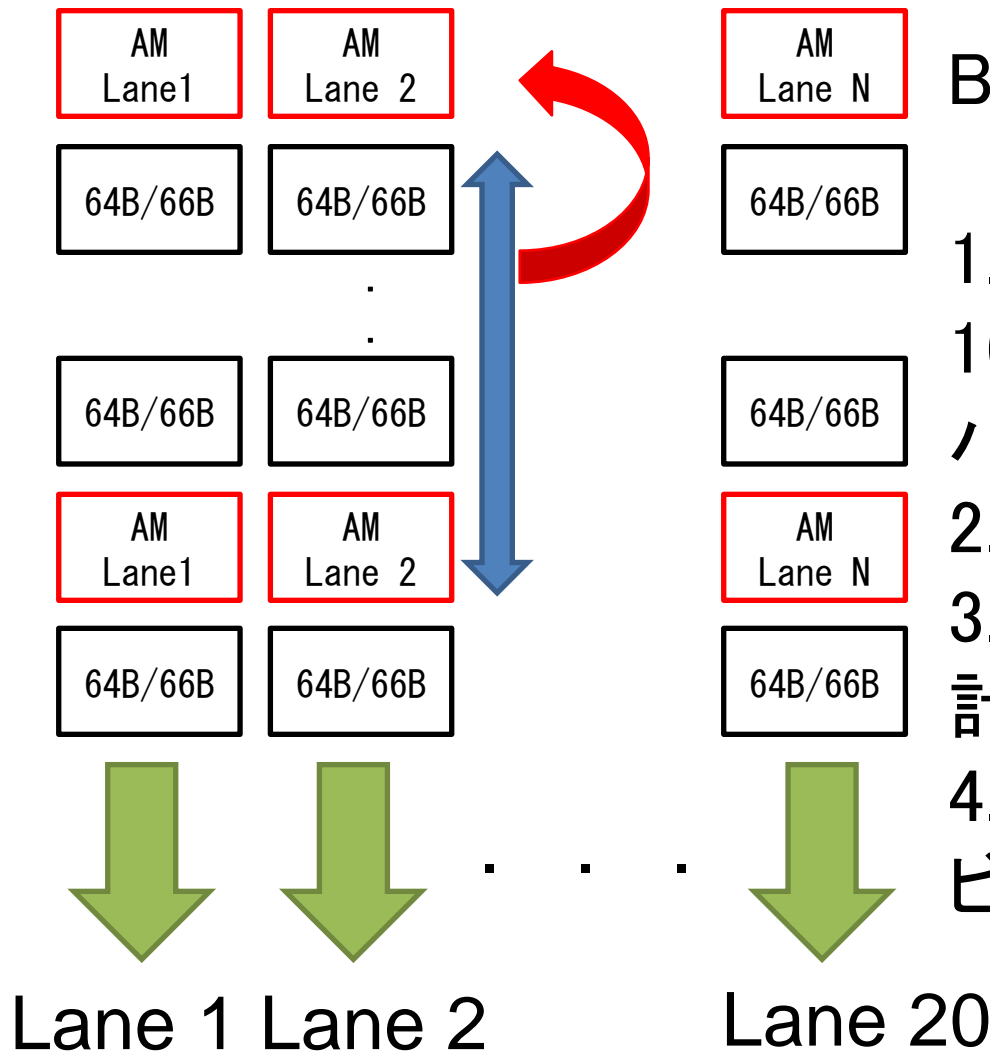


TX側とRX側のレーン対応関係を伝えるため、データ送信を一時やめてアライメントマーカ―を全レーンに同時に挿入

AMの役割

- RX側でレーン対応関係の判別
- レーン毎の到着時間のズレ (Skew)の吸収

BIP(Bit Interleaved Parity)



BIPの仕組み概要

1. 送信側でレーン単位で16383+1ブロックのデータのパリティを計算
2. AMにパリティを埋め送信
3. 受信側で同様にパリティ計算し、AMのパリティと比較
4. 異なればそのレーンでビットエラー発生をカウント

BIPカウンター表示機能について

•実装状況(試験時点で)

- Brocade → 実装されていない
- Cisco → 実装されている
- Juniper → 詳細はJuniper社の発表にて

BIPカウンターの表示例 (Cisco)

```
RP/0/RP0/CPU0:cisco_CRS#sh controller hundredGigE 0/1/0/0 phy
```

```
PCS Lane BIP Error Counters:
```

Lane-0 : 0	Lane-10 : 0
Lane-1 : 31	Lane-11 : 0
Lane-2 : 0	Lane-12 : 0
Lane-3 : 0	Lane-13 : 39
Lane-4 : 52	Lane-14 : 0
Lane-5 : 0	Lane-15 : 0
Lane-6 : 0	Lane-16 : 0
Lane-7 : 0	Lane-17 : 46
Lane-8 : 39	Lane-18 : 0
Lane-9 : 0	Lane-19 : 0



- 可変アッテネータを入れ信号を劣化させた場合の試験
- 5レーンでBIPエラーが発生
- 1波長で正常に受信が出来なくなっている可能性を示す

```
Total PCS Lane BIP Error Count : 207  
Total PCS Lane Sync Header Error Count : 13  
Total PCS Lane Bad 64/66 Code Count : 204
```

100GBASE-LR4は非常にセンシティブ

•実験の際に確認された現象

- 100GE I/FはUP。しかし、対向とのpingロスが多発
- I/Fエラーカウンター、送受信光レベルに問題はなし
→ 光ファイバーの端面清掃 + 光SW+CFPの光素子部分の清掃で解決

•想定される原因

- 光が強すぎたので5dBのアッテネータを挿入してた
- CFPの問題(光素子がセンシティブ、WDM)
- 100GEのPCS層の不具合

•懸念点

- 構内接続であれば磨けばなんとかなりそう
- ビル間接続のダークファイバー利用時は不安が…

再度、外山さんへ