



BroadBand Tower, Inc.

JANOG44 発表資料

# データセンター間接続（DCI）を 1から作り直す作り直している話 ～実践編～

---

2019年07月26日

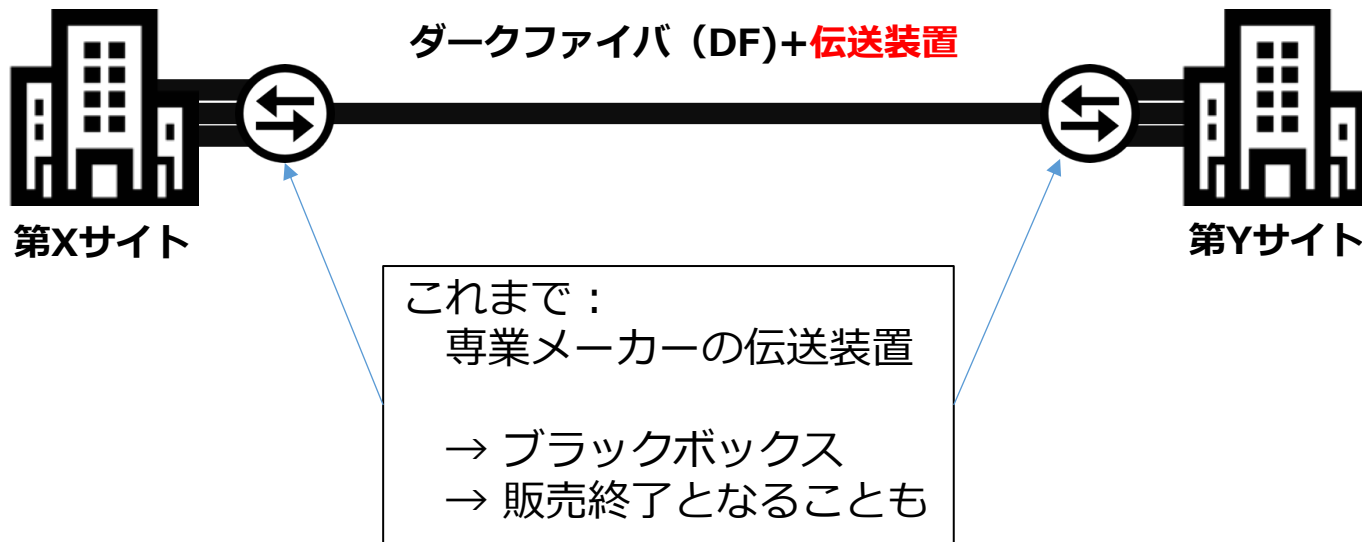
株式会社ブロードバンドタワー  
Cloud&SDN研究所  
西野大

## データセンター間接続（DCI）を一から作り直す話

- 動機&ゴール：           なぜ作り直そうと思ったのか？
- 第1ステップ：           ラボ内試験
- 第2ステップ：           データセンター内接続
- 第3ステップ：           データセンター間接続
- ネクストステップ：   100Gbpsオーバーの世界

## データセンター間接続（DCI）を一から作り直す話

- 動機&ゴール：           なぜ作り直そうと思ったのか？
- 第1ステップ：           ラボ内試験
- 第2ステップ：           データセンター内接続
- 第3ステップ：           データセンター間接続
- ネクストステップ：   100Gbpsオーバーの世界



## ブロードコムが Brocadeを買収

Top > サーバ / ストレージ / ネットワーク

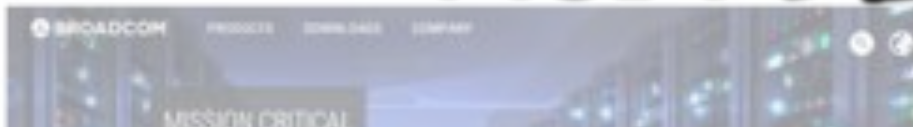
半導体ベンダーのブロードコムがブロードコムの買収を発表。  
ただし欲しいのはFC SAN事業だけ、イーサネット事業は買収後すぐに手放すと

### カテゴリ

- ・ クラウド
- ・ Docker / コンテナ / 仮想化
- ・ 開発ツール / 言語 / プログラミング
- ・ DevOps / アジャイル開発
- ・ システム / システムインテグレーション
- ・ Windows / Linux
- ・ ネットワーク / ストレージ / ネットワーク
- ・ NoSQL / モデルウェア
- ・ AI / ビッグデータ
- ・ アプリケーション / Office
- ・ 働き方 / 給与 / 学び
- ・ 業界動向 / IoT / その他

# ネットワーク製品ベンダーの M&Aの衝撃

通信機器関連の半導体ソリューションベンダー大手ブロードコムは、ストレージ、ネットワークやイーサネットスイッチ製品、データセンター向けサーバーシステムズの買収を発表しました。



参照 [https://www.publickey1.jp/blog/16/fc\\_san.html](https://www.publickey1.jp/blog/16/fc_san.html)

# 2018年：Infinera が Coriant を買収



The screenshot shows the Photonics Japan website. The main navigation bar includes categories like PHOTONICS, BIO PHOTONICS, INDUSTRIAL LASER, LED, and MACHINE VISION. A prominent banner for LaserFocusWorld Japan features Lumentum products. Below this is a navigation menu with items like TDP, ニュースアーカイブ, 国内レポート, etc. A secondary banner for LDMblue and laserline is visible. The main content area shows a breadcrumb trail: TOP > ニュース > Business/Market > Infinera, 4億3000万ドルでCoriantを買収. The article title "Infinera, 4億3000万ドルでCoriantを買収" is circled in pink. The article text, dated July 27, 2018, from Sunnyvale, CA, reports that Infinera has announced the acquisition of Coriant, a move that significantly expands its scale. It details how this acquisition allows Infinera to serve as a full-service provider for network operators, moving from a pure-play equipment vendor to a more integrated solution provider. The article also mentions that this move aligns with industry forecasts from Ovum, which predicts a 178 billion dollar investment in optical networks by 2022.

参照 <http://ex-press.jp/lfwj/lfwj-news/lfwj-biz-market/25162/>

- ✓ M&Aのニュースは、予徴なく突然やってくる
- ✓ 「これは！」と思う「キラリ」とした  
「導入したいと思う製品」を作っている会社が、M&Aされる
- ✓ 「製品ラインアップは維持します」と公式見解はでることは多いが・・・、  
突然に消える製品群
- ✓ 最近は「地政学的なリスク」も考えないとならない

光伝送装置では一部だけ取り替えがきかないことが多く、インパクト大

- ✓ M&Aのニュースは、予徴なく突然やってくる
- ✓ 「これは！」と思う「キラリ」とした  
「導入したいと思う製品群」がM&Aで消える  
**対応策の一つとして**
- ✓ 「製品ラインアップは維持します」と公式見解はでることは多いが・・・、  
突然に消える製品群

## Open Optical Line System

光伝送装置では一部だけ取り替えがきかないことが多く、インパクト大



## ■ Telecom Infra Project (TIP)

- ◆ Facebookが提唱（2016年～）、現在、数百社が参画
- ◆ きっかけ： Open Compute Project の成功をインフラに展開
- ◆ ねらい： 通信事業者のインフラ向け技術のオープン化
- ◆ ターゲット： アクセス、バックホール（D-WDMなど）、コア



## ■ Open Optical Packet Transport

- ◆ TIPの一つのプロジェクトグループ（PG）
- ◆ 「通信事業者のバックホール向けのD-WDMなどの光伝送技術を**オープン化**」が目標
- ◆ **Open Optical Line System**という考え方を提唱
- ◆ 光伝送技術でも、**disaggregation** を推進

嵐の海をこの先生きのこるには？

できるだけオープンな標準的技術を使おう

光伝送装置のdisaggregationに挑戦しよう

# 東京都区内の近距離データセンター間



## 1Gbps・10Gbpsのイーサネット

## ■ 中継ナシ

- 「減衰」がエンドエンド10~30dBなので、ダークファイバの中間で増幅しなくてもよい。
- インラインの**中継光アンプ**の**設備設置**を考慮しなくて良い。

## ■ 分散ムシ

- 速度が10Gbps以下、距離が80km以下、光ファイバの「分散」の影響を考慮しなくて済む。
- **分散補償**などのデバイスが**不要**。

## ■ Cバンド（1550nm近傍）利用

- CバンドのITU-Tグリッド対応のD-WDMの機材（MUX/DEMUX、SFP・SFP+）は、入手が容易。
- Cバンドだと、安価な**EDFA**タイプの光アンプを利用可。（EDF = Erbium Doped Fiber Amplifier）

## ■ 100Gbpsは将来・・・

- 100Gbpsでは、CバンドでのカジュアルなOpticsが存在しなかった。

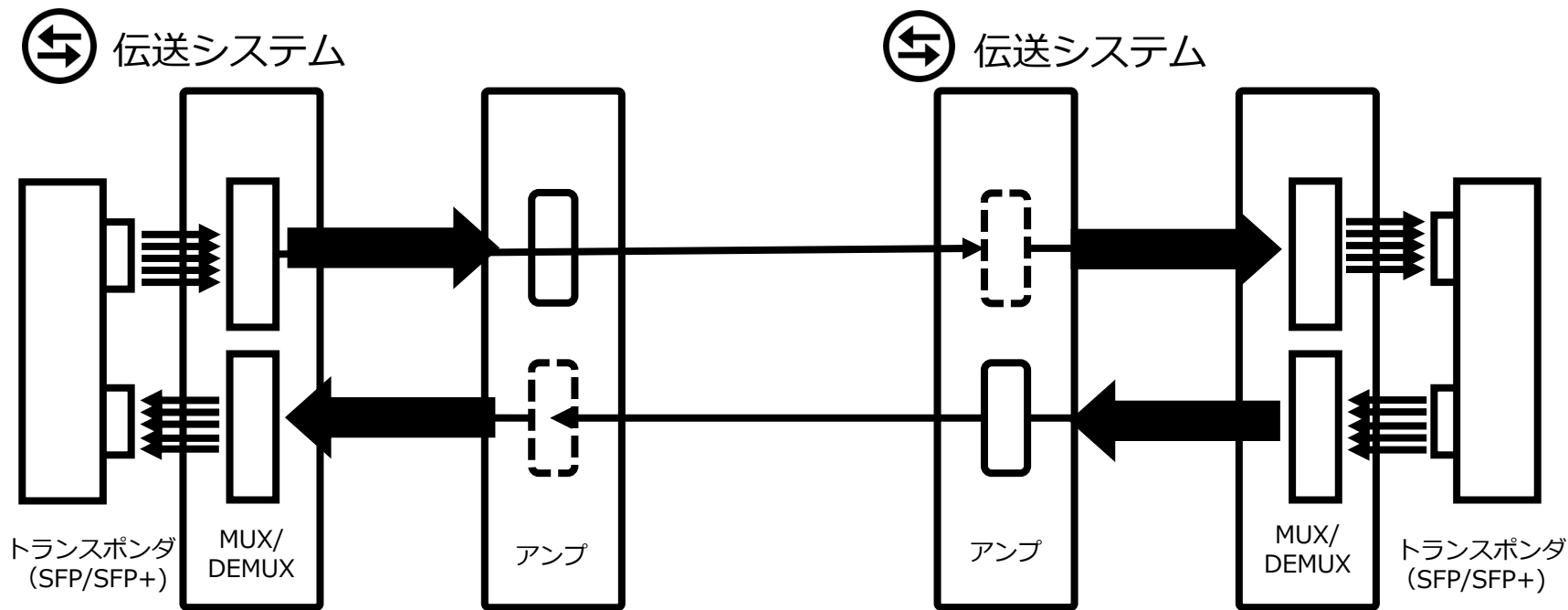
## ■ そもそも・・・

- 研究所でPoCテストに利用できるのが、23区内のデータセンターだった。

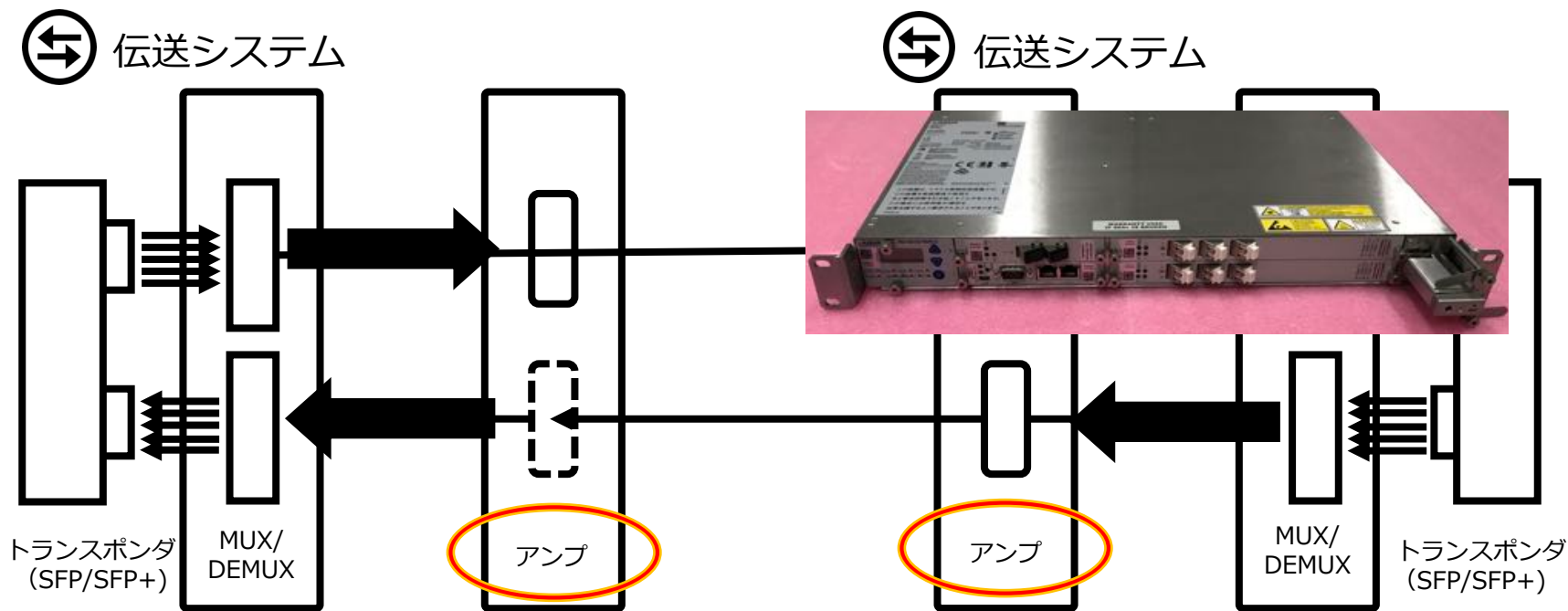
## データセンター間接続（DCI）を一から作り直す話

- 動機&ゴール：           なぜ作り直そうと思ったのか？
- 第1ステップ：           ラボ内試験
- 第2ステップ：           データセンター内接続
- 第3ステップ：           データセンター間接続
- ネクストステップ：   100Gbpsオーバーの世界

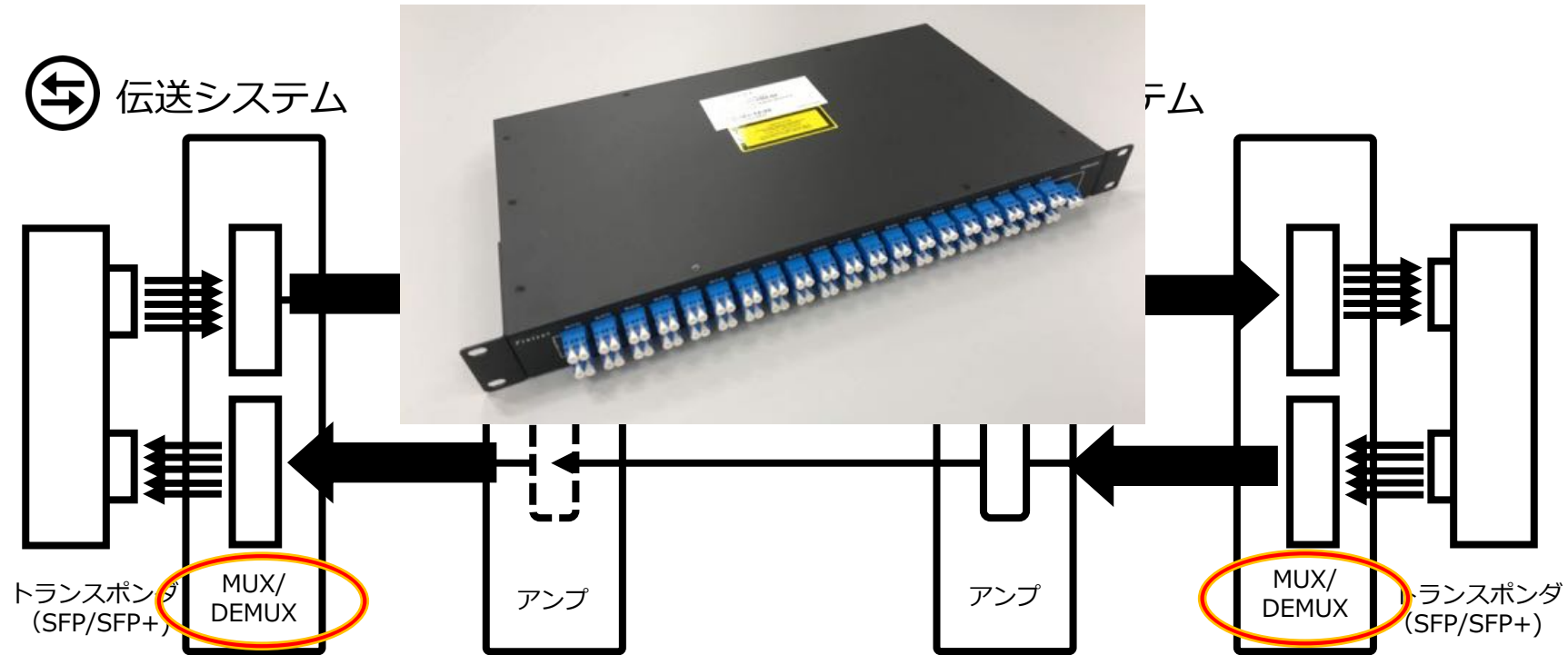
# 今回のシステム



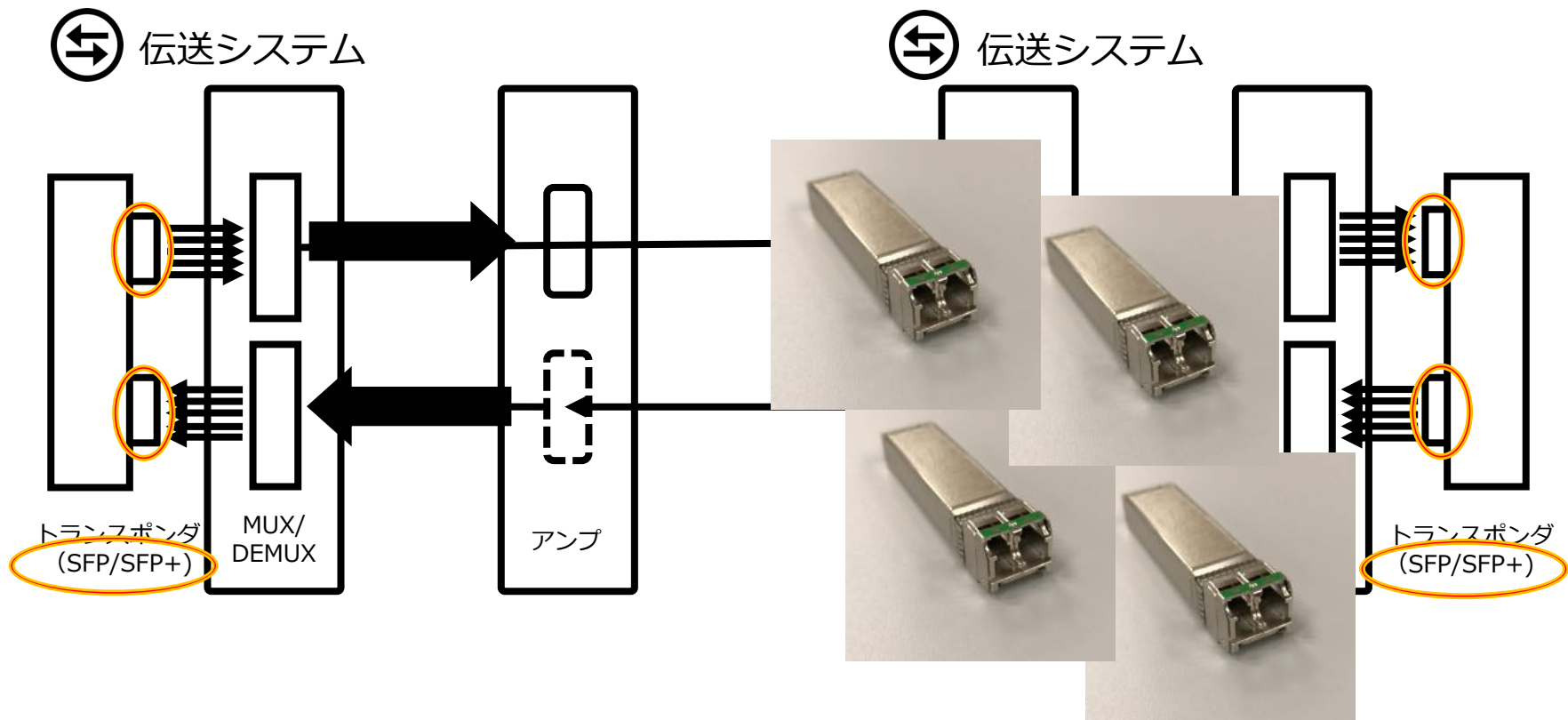
# 必要なパーツ：アンプ（増幅器）



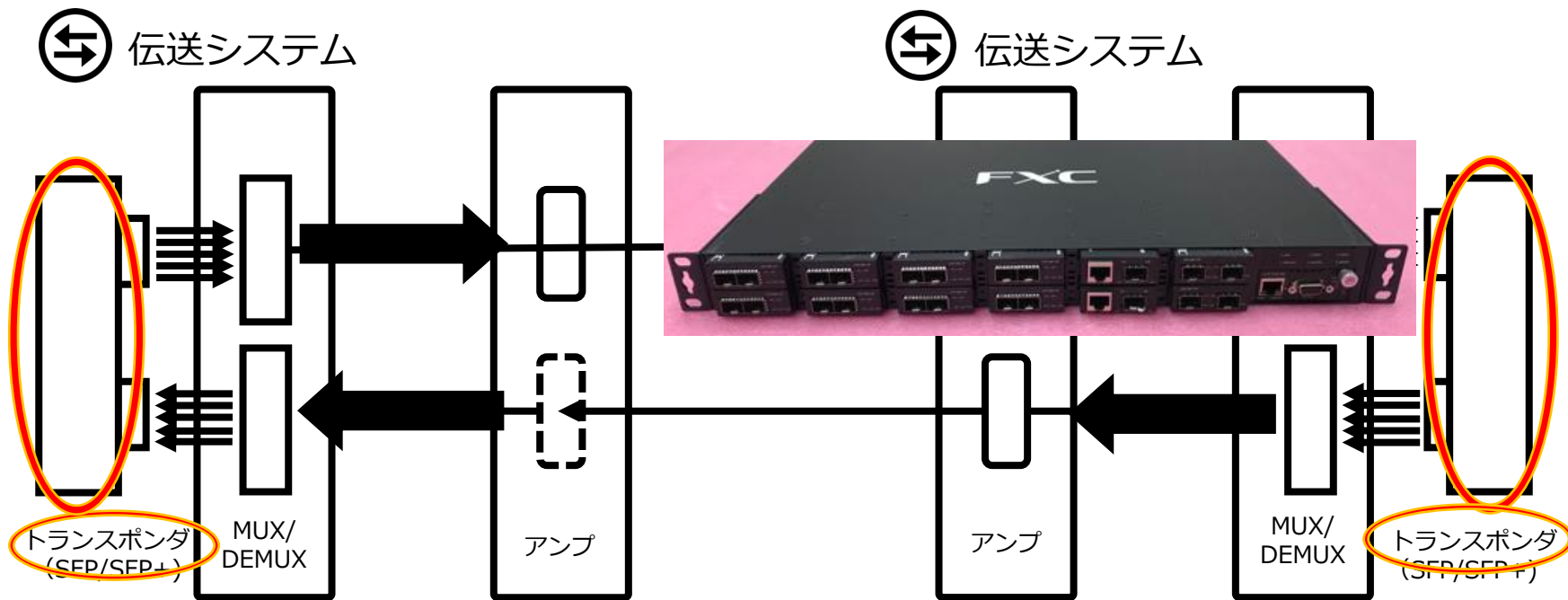
# 必要なパーツ：MUX/DEMUX（合波分波器）







# 必要なパーツ：コンバータ（トランスポンダ）

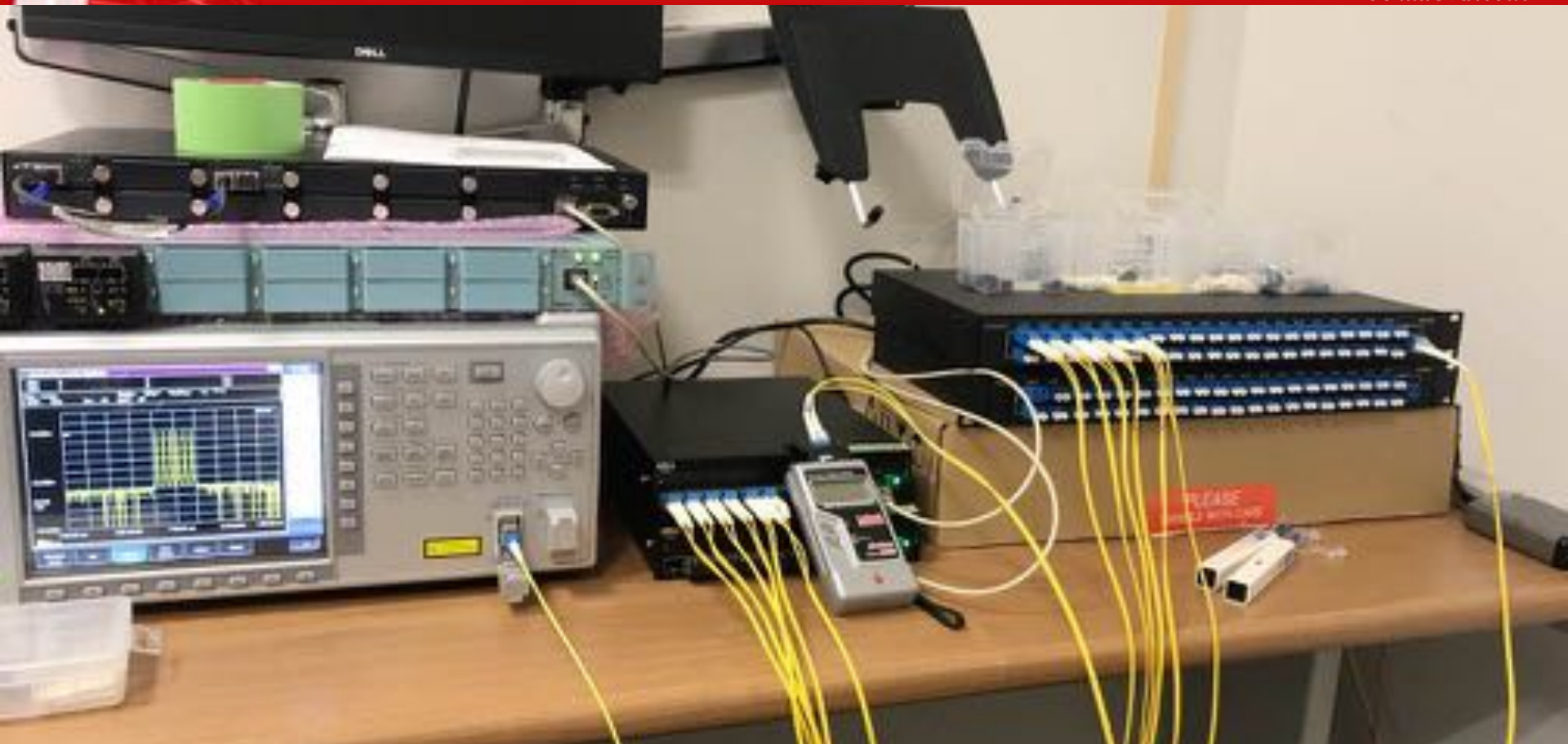


④ 伝送システム 装備 (パーツ) を買って・・・

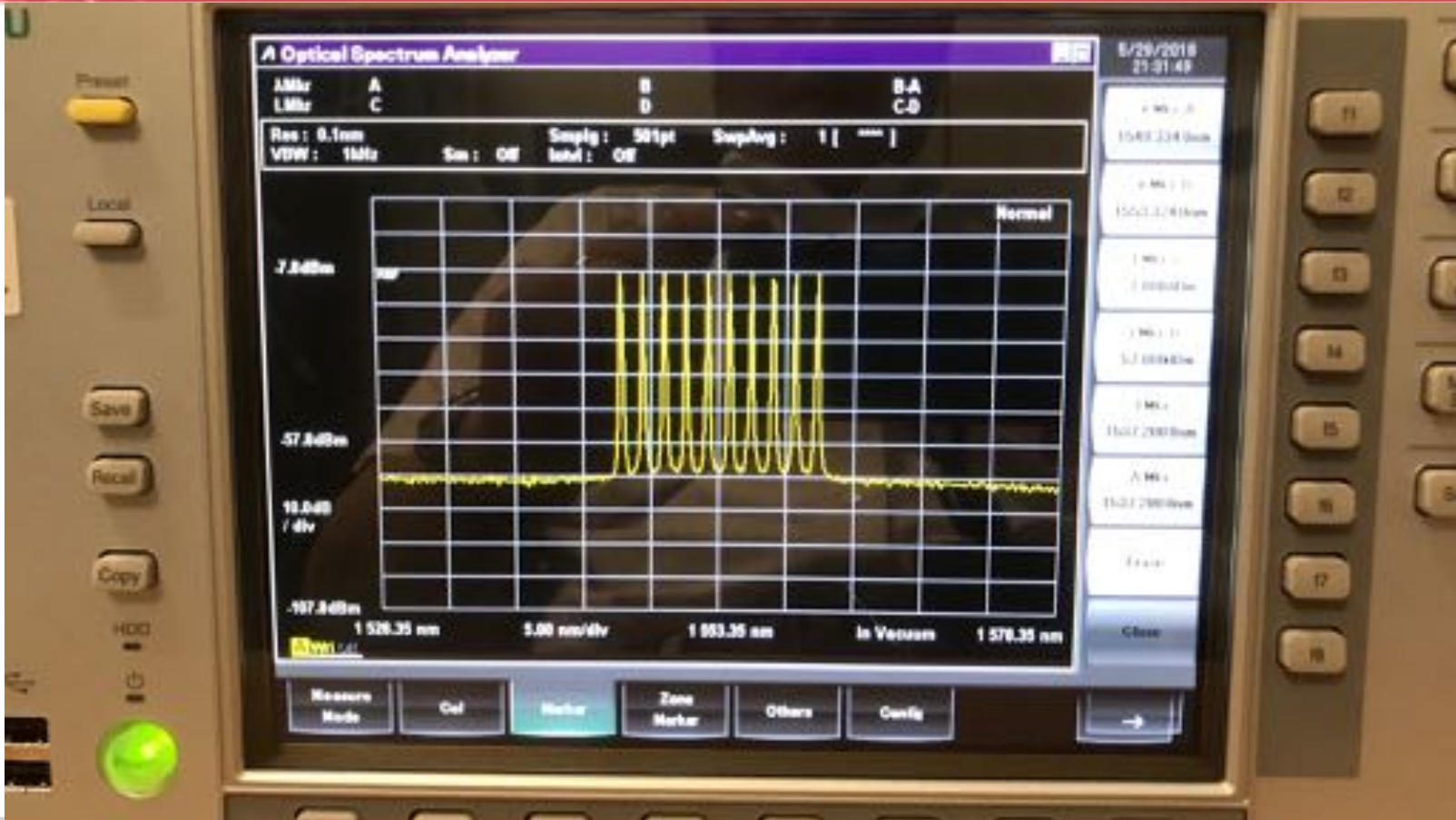
# ラボでPoC試験を実施



※PoC=Proof of Concept



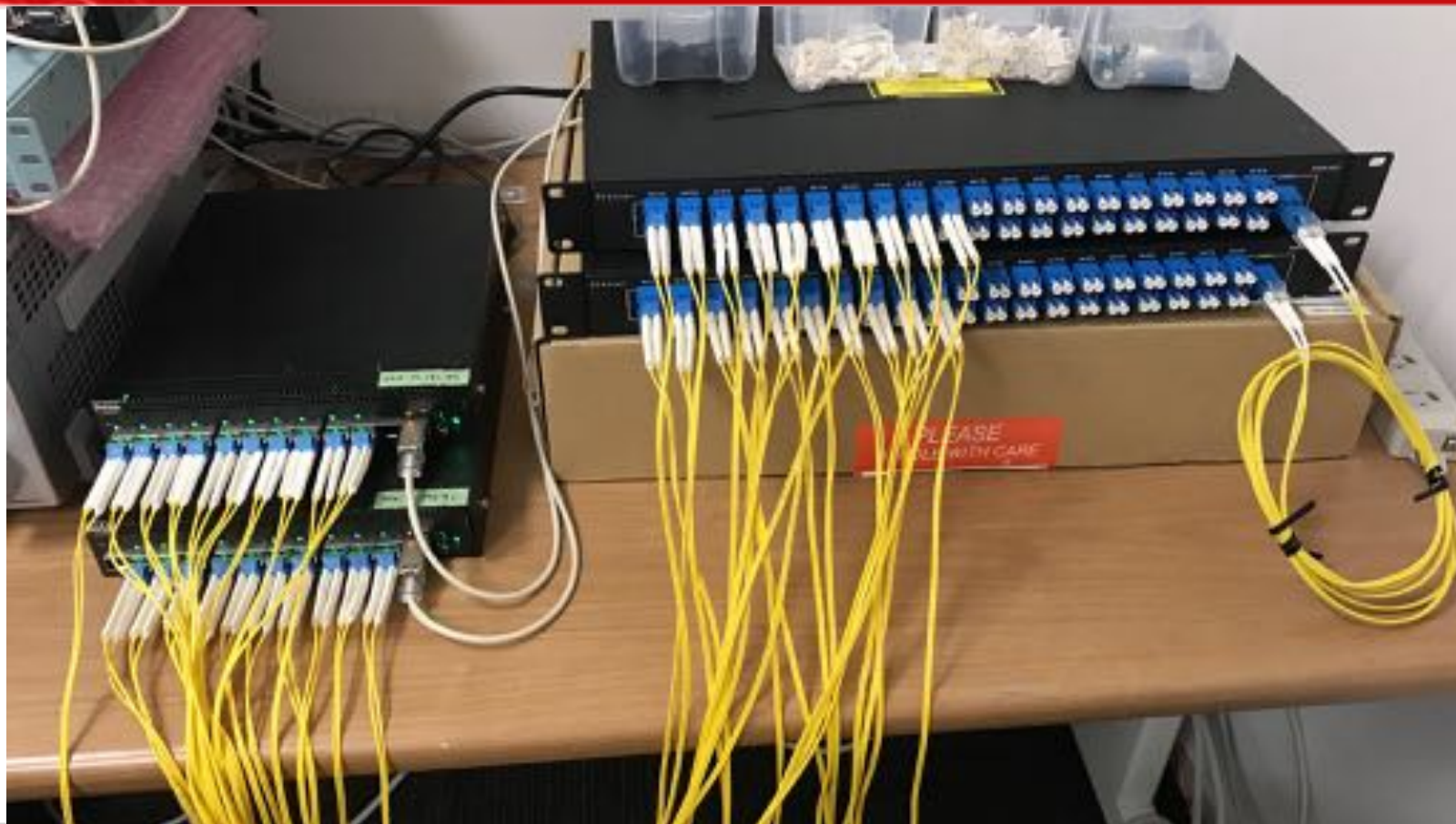
# ちゃんと10波ピークがみえる







# 当社のリングホール





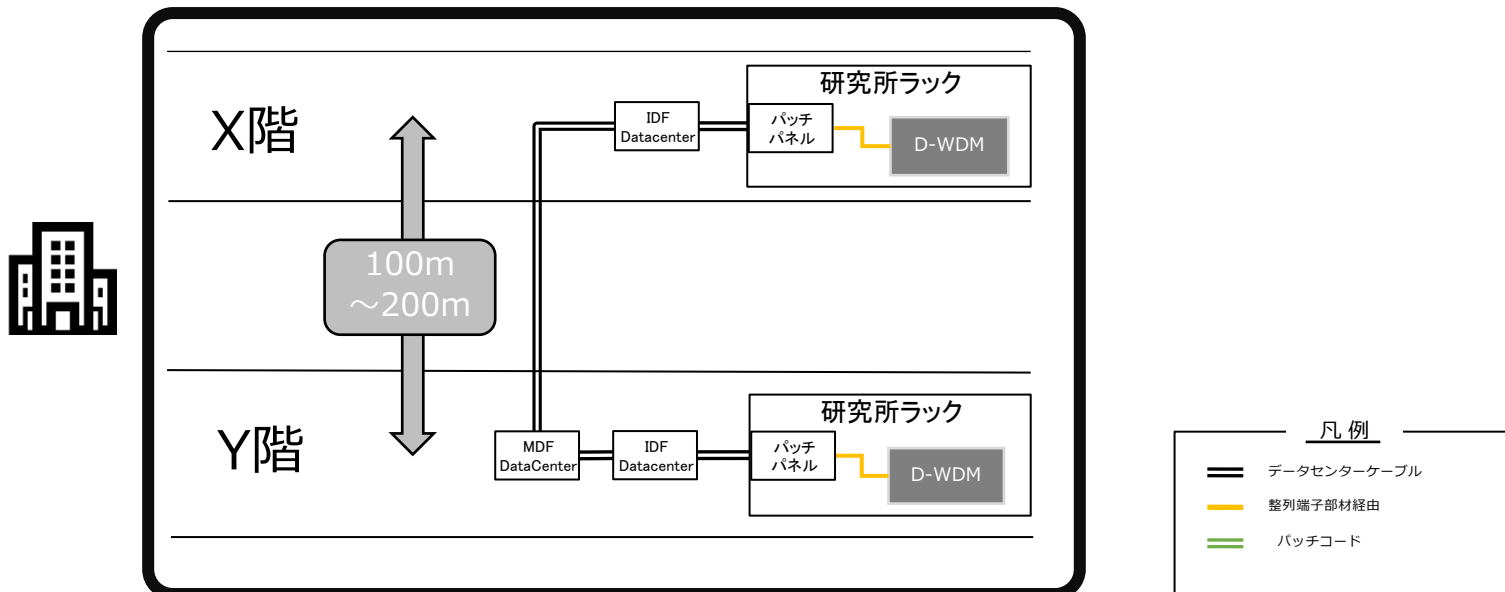


# 0.001マイルの長距離？伝送に成功

## データセンター間接続（DCI）を一から作り直す話

- 動機&ゴール： なぜ作り直そうと思ったのか？
- 第1ステップ： ラボ内試験
- **第2ステップ： データセンター内接続**
- 第3ステップ： データセンター間接続
- ネクストステップ： 100Gbpsオーバーの世界

## データセンター内でDCIの技術（D-WDM）を適用する



一般のデータセンターで複数フロアを使っているイメージ

## データセンター内の階間光ファイバ



残り 1 対

かなりの芯数を敷設していても  
何故か？足りなくなる??

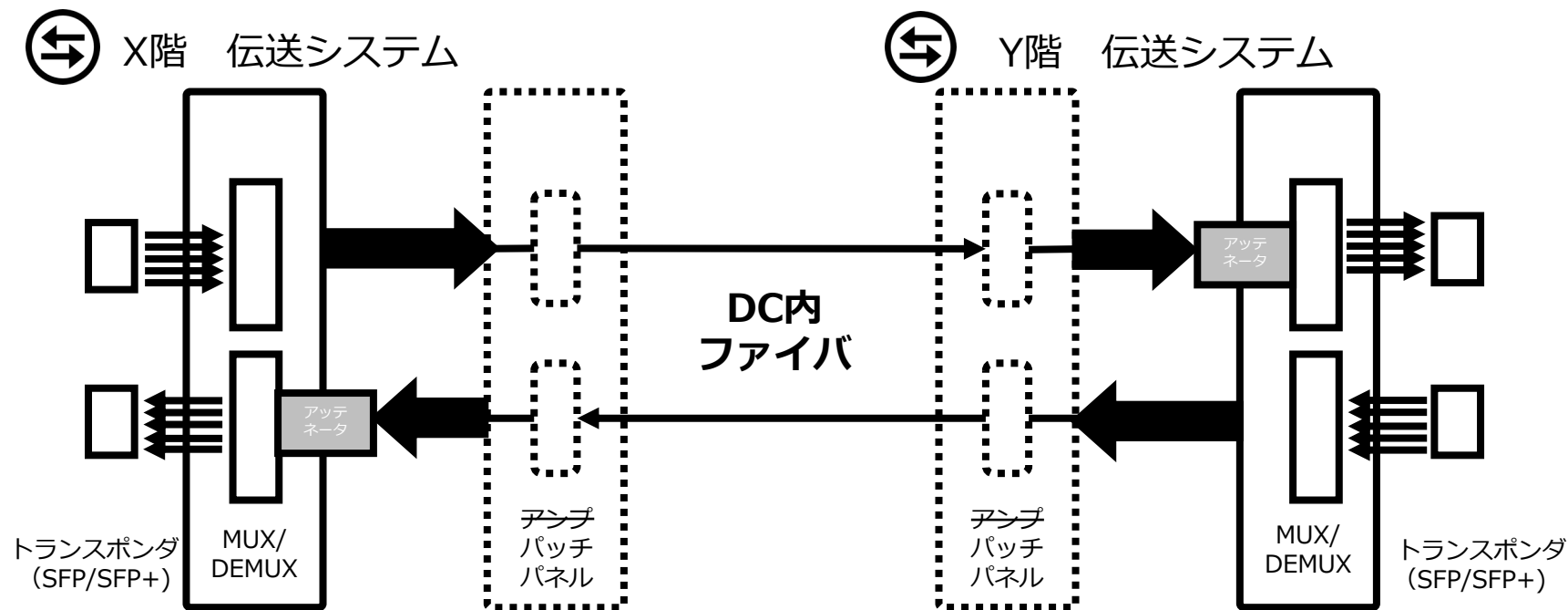
データセンターにおいて  
配線の追加リクエストは  
コストが大きい（時間、\$\$\$、……）



「配線を追加工事するくらいなら

データセンター内でD-WDMを！」

# データセンター内接続：構成図



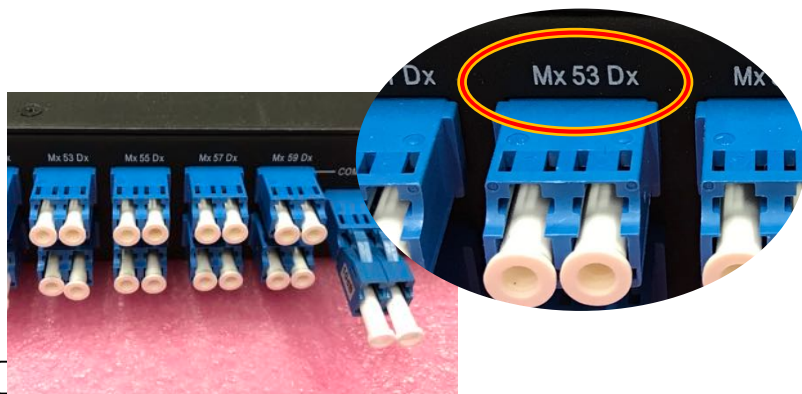
DC内ファイバは損失が非常に少ない → 5dB~10dBのアッテネータが必要

## ■ MUX/DEMUX と Optics (SFP/SFP+)

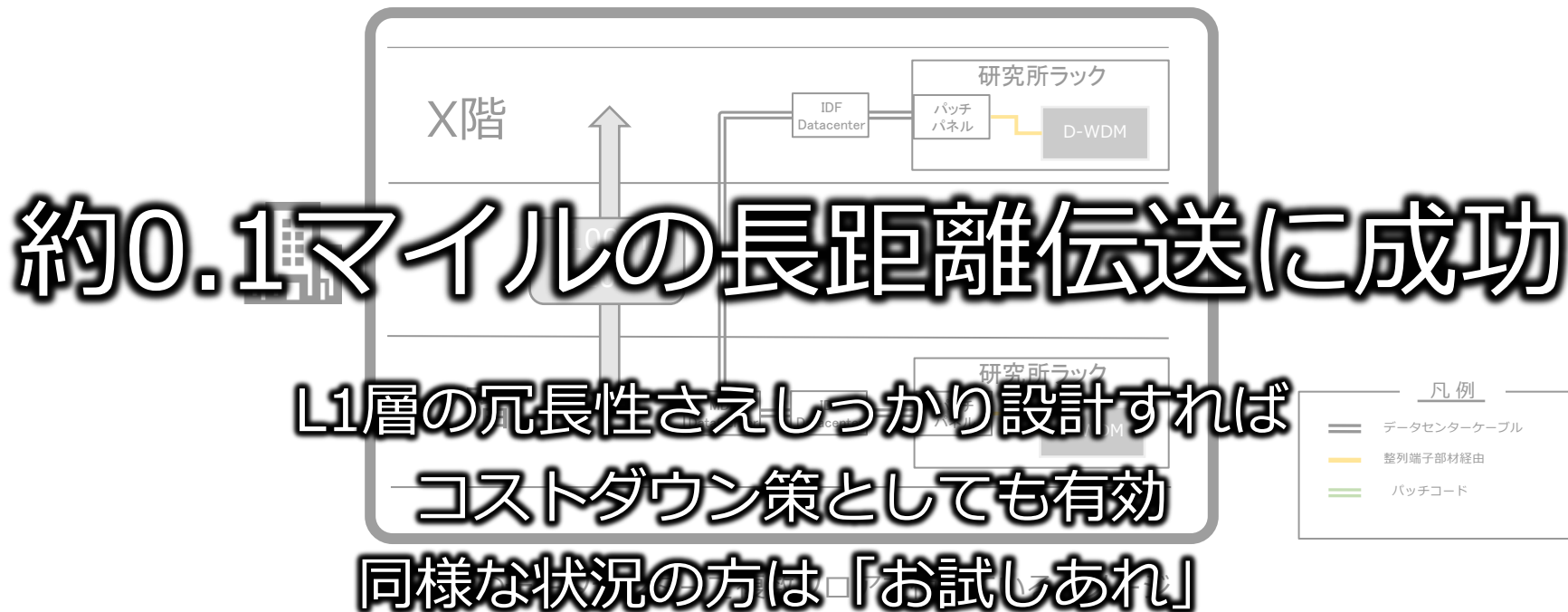
- ◆ 1G/10Gbpsなら、ITU-T G.694.1 100GHzグリッドの製品がポピュラー
- ◆ 固定波長タイプなら、「チャンネル番号」を合わせれば結線するだけ = スイッチの感覚

## ■ 注：チャンネル番号

- ◆ チャンネル#  $\Delta\square = 19\Delta\square00$  GHz
- ◆ チャンネル番号は、上記の表現がデファクトだが、メーカー独自のケースもあるので注意



データセンター内でDCIの技術（D-WDM）を適用する



## データセンター間接続（DCI）を一から作り直す話

- 動機&ゴール：           なぜ作り直そうと思ったのか？
- 第1ステップ：           ラボ内試験
- 第2ステップ：           データセンター内接続
- 第3ステップ：           データセンター間接続
- ネクストステップ：   100Gbpsオーバーの世界

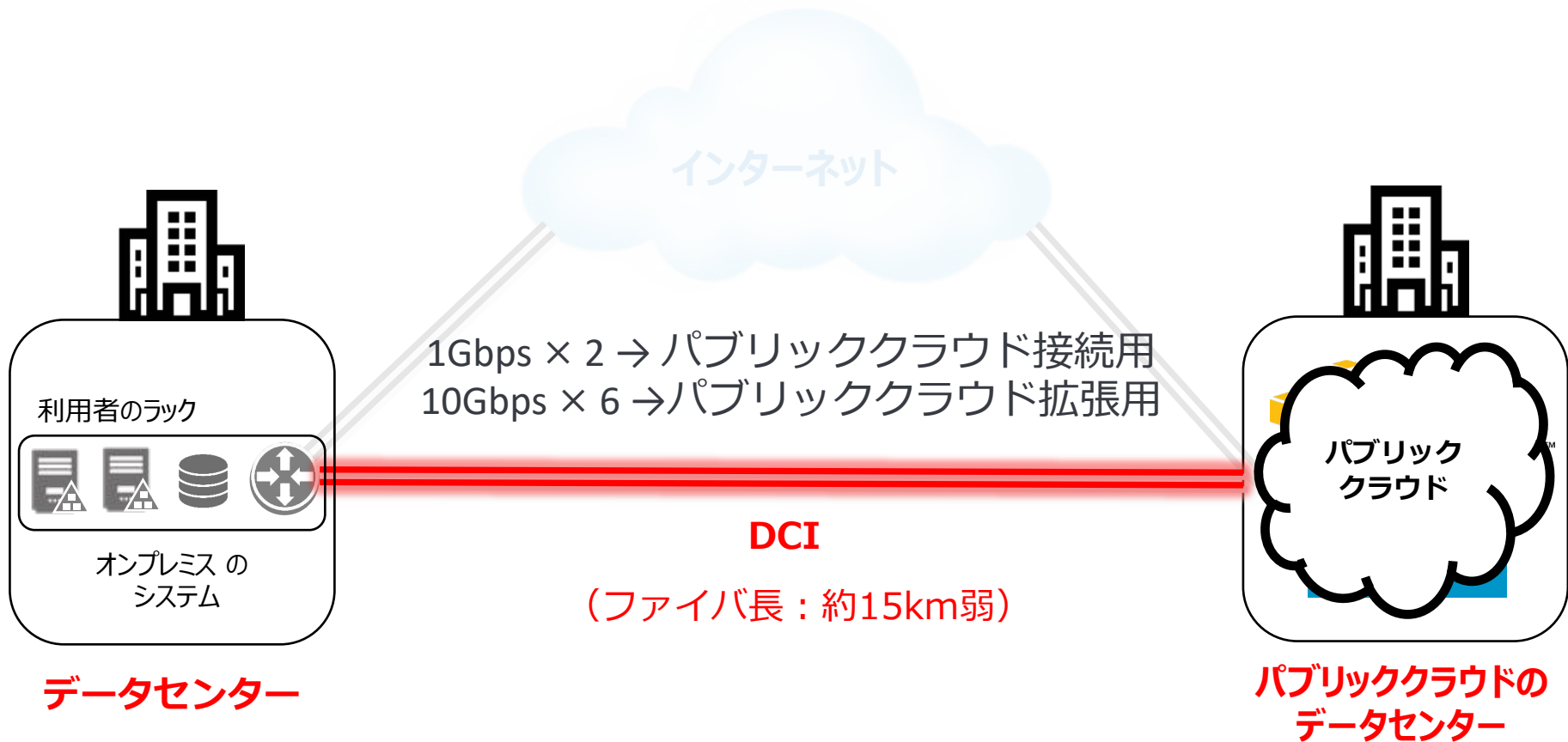


データセンター内接続で経験値をアップした後は

# 次は町へ冒険にでよう

…ようやく**光アンブ**が必要な世界へ

# 適用するネットワーク



disaggregation の冒険には

# 装備が重要

ガジェット・測定器・・・

## ■ 必須なモノ

- ◆ ファイバ・メンテナンス用品（クリーナー）
- ◆ パワーメータ・・・（+25dBmあたりまでを測定できるもの）
- ◆ アッテネータ
- ◆ 光パッチコード・・・（1m程度の短いものを大人買い）
- ◆ **防護メガネ**

## ■ できれば持つべきモノ

- ◆ **スペクトラム・アナライザー**
- ◆ OTDR

## ■ Why ?

- ◆ D-WDMでは、MUX/DEMUXや光暗譜の出力の光レベルは非常に強くて危険
  - 例：EDFAアンプ 出力 最大値                    ~ +27dBm
  - 例：SFP+ (0dBm出力) × 40波 MUX出力   ~ +16dBm
- ◆ 「出力光をのぞき込まない」は当たり前！！

「光出力端子にキャップがつけ忘れた装置がある」「ファイバの先を向けてしまう」などの万が一のミスはあり得るので防護メガネを着用。

## ■ 選び方のコツ

- ◆ フィルタ波長：
  - 1250nm~1650nm の赤外光に対する光学濃度が十分に大きい
- ◆ 可視光：
  - フィルタ波長の光学濃度は大きい方が良いが、可視光の光学濃度は小さくないと、作業が不便
- ◆ その他
  - オーバーグラス型だと、万人に嬉しい

# ガジェット：レーザー防護メガネの例



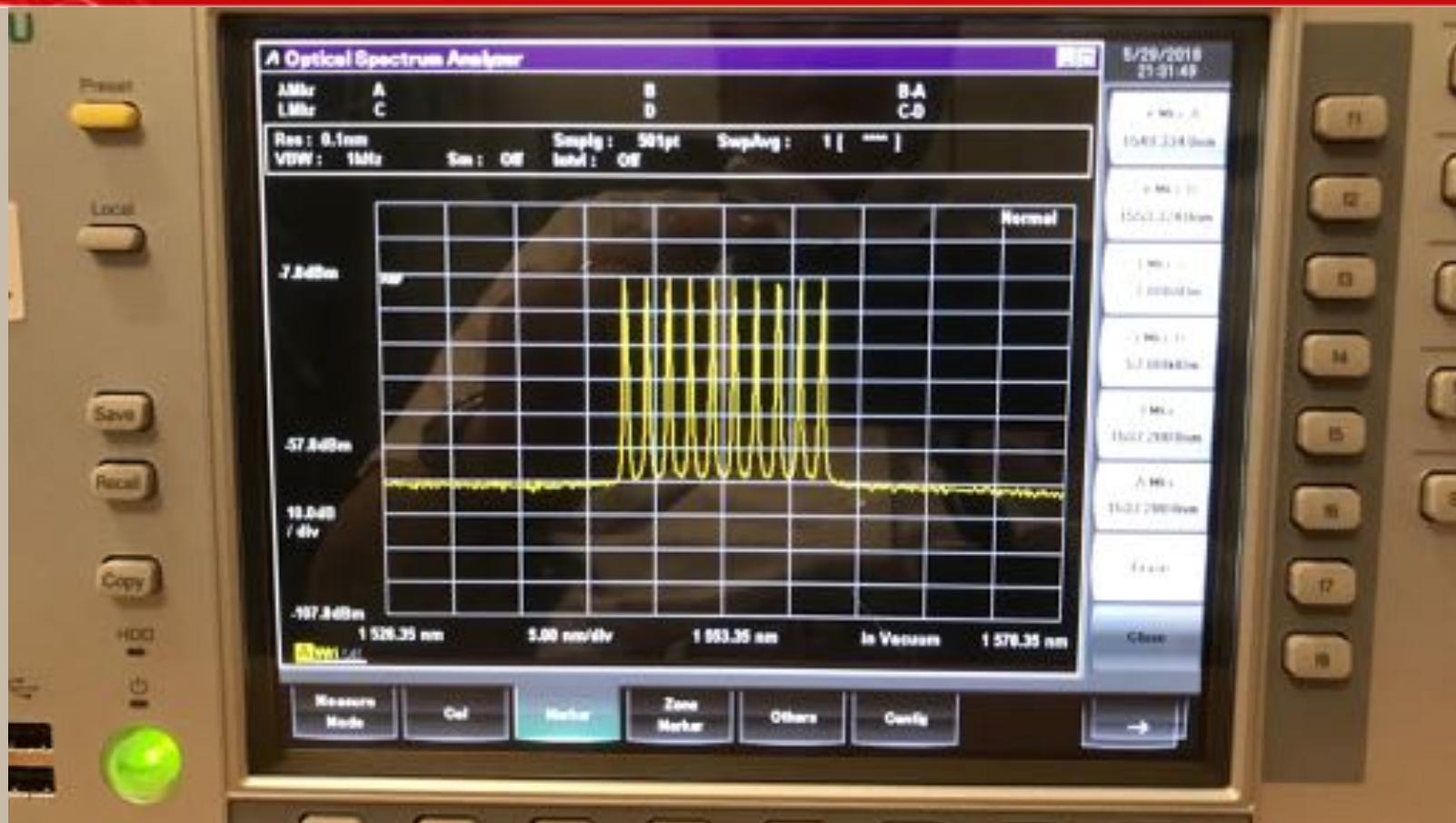
理研オプテック  
RS-80 YG

山本光学  
YL-250 GヤグOD5

# 据置き型スペクトルアナライザ

アンリツ株式会社 MS9740A

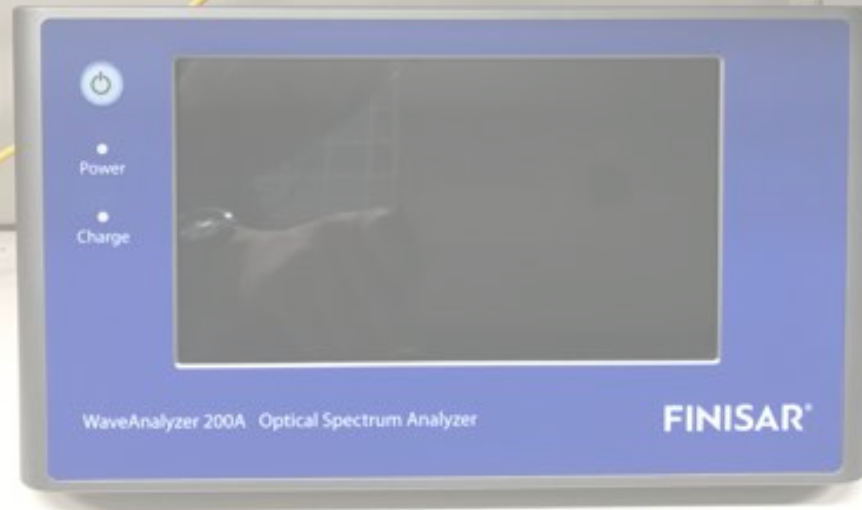


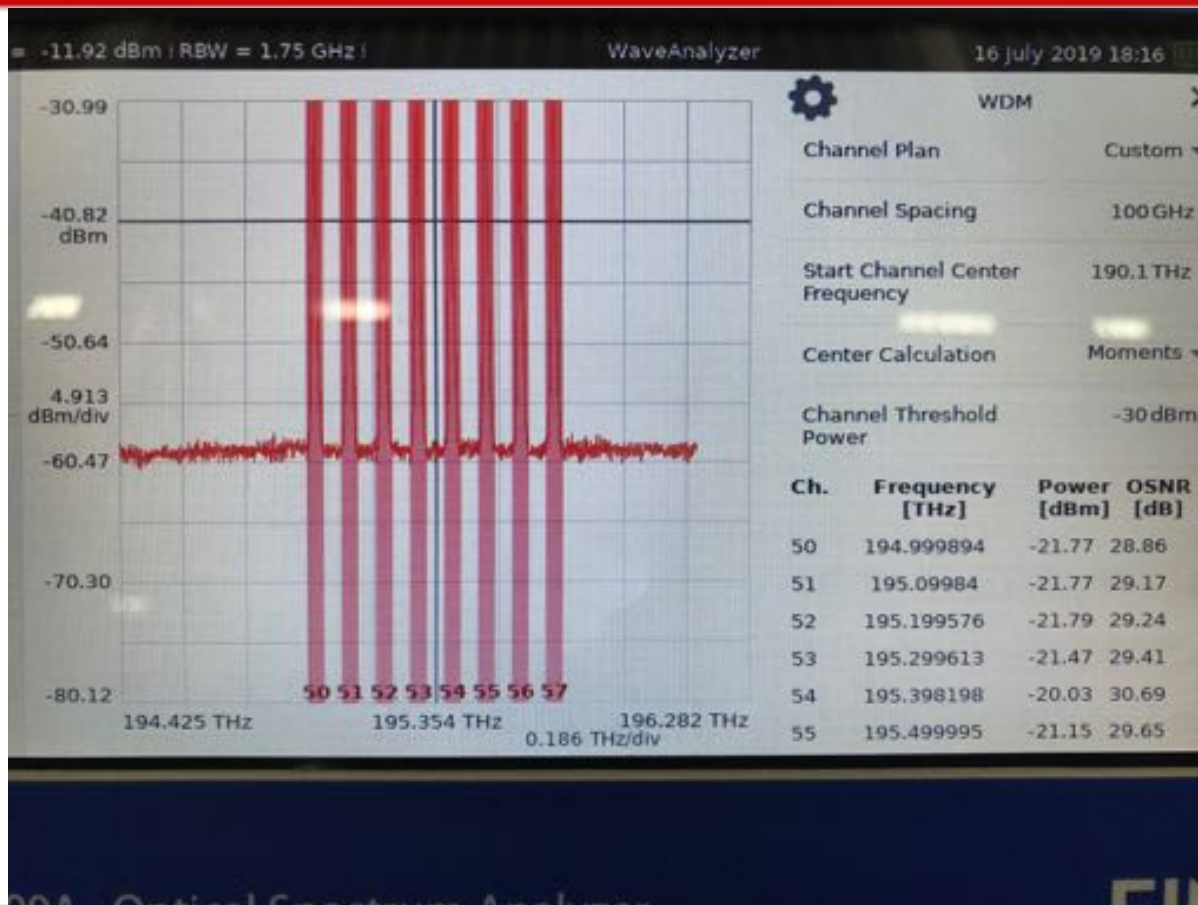




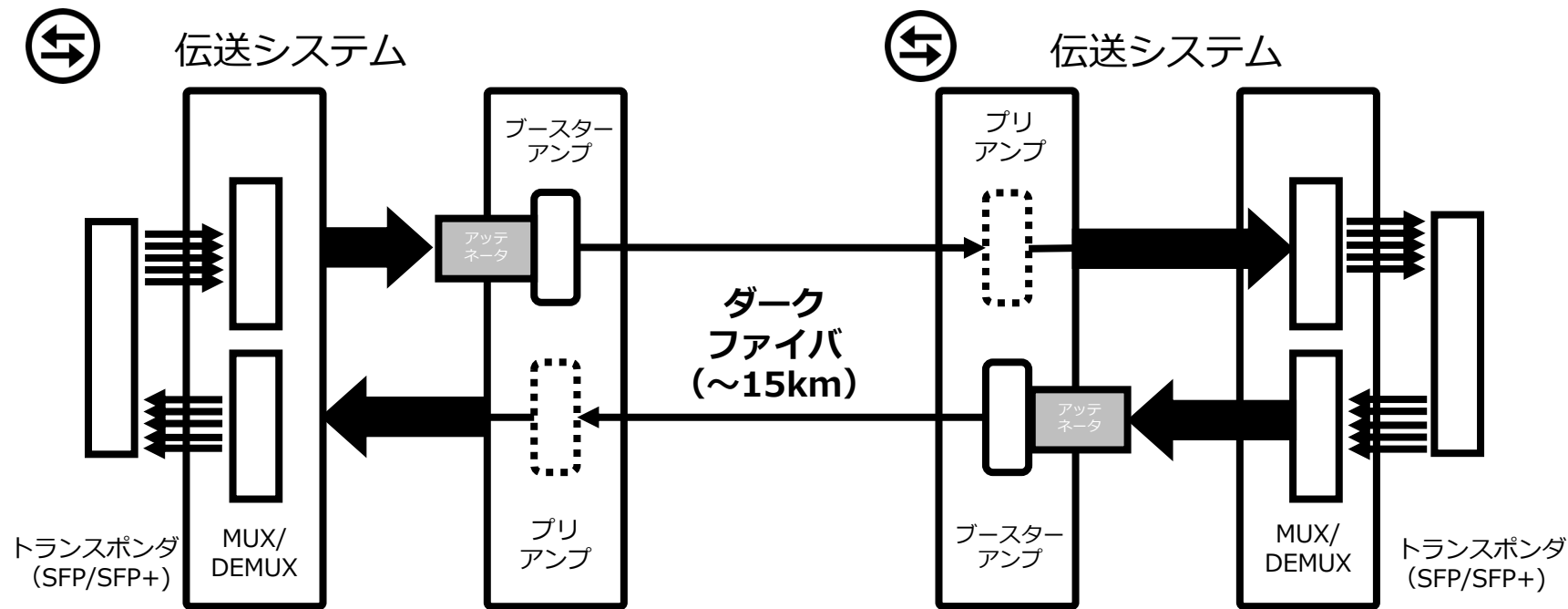
# 携帯型スペクトルアナライザ

FINISAR社 WaveAnalyzer 200A





# データセンター間接続：構成図



ダークファイバの減衰が小 → 受信側のプリアンプ不要、ブースターアンプのみ

## アンプの入力上限値を超えないように調整を行う

### ■ ブースターアンプの入力上限値

- ◆ MUX で合波した光は、かなり強い

- 例：-1dBmの光でも、8波合成すると、約+9dBされる（合計約8dBmとなる）

- ◆ アンプの入力許容値を超えないように注意

→ スペックシートを確認 → アッテネータ（固定 or 可変）で調整

## Opticsの入力上限値を超えないように調整を行う

### ■ 受光側Opticsの上限値

- ◆ 長距離用のトランスポンダを直結すると、入力許容値を超える場合がある
- ◆ 終端するOpticsの入力許容値を超えないようにする必要がある
  - ブースタアンプ or プリアンプ で全体の増幅値を調整
  - Optics直前にアッテネータ（固定 or VOA）を入れて調整

## IV. Optical Characteristics (TOP = -5 to 70 °C, VCC = 3.14 to 3.46 Volts)

Parameter	Symbol	Min	Typ	Max	Unit	Ref.
<b>Transmitter (Tx)</b>						
Average Launch Power	$P_{AVE}$	-1		+3	dBm	
Optical Wavelength	$\lambda_c$	As per ITU-T 694.1			nm	1
Center Wavelength (End of Life)	$\lambda_{c\_EOL}$	$\lambda_c \pm 100\text{pm}$				
Side-Mode Suppression Ratio	SMSR	30			dB	
Optical Extinction Ratio	ER	8.2			dB	
Average Launch power when Tx is OFF	$P_{OFF}$			-30	dBm	
Relative Intensity Noise	RIN			-128	dB/Hz	
<b>Receiver (Rx)</b>						
Optical Center Wavelength	$\lambda_c$	1260		1600	nm	
Reflectance	$R_{rx}$			-27	dB	
Sensitivity (0km)	Bit Rate (Gb/s)	BER	$R_{SENS1}$	-24	dBm	2
	8.5, 9.95-10.7	$<10^{-12}$				
Sensitivity (80km)	8.5, 9.95-10.7	$<10^{-12}$	$R_{SENS2}$	-21	dBm	2.3.a
			$R_{SENS3}$			2.3.b
			$R_{SENS4}$			2.3.c
Overload (Average Power)	$P_{AVE}$	-7			dBm	
LOS Deassert	$LOS_D$			-40	dBm	
LOS Assert	$LOS_A$	-37		-30	dBm	
LOS Hysteresis	$LOS_H$	0.5			dB	

最大入力・最小入力

## II. Absolute Maximum Ratings

Exceeding the limits below may damage the transceiver module permanently.

Parameter	Symbol	Min	Typ	Max	Unit	Ref.
Maximum Supply Voltage	Vcc	-0.5		4.0	V	
Storage Temperature	$T_s$	-40		85	°C	
Receiver Optical Damage Threshold	RxDamage	5			dBm	1

これを超えてはイケナイ

## アンプの入力光を適性に揃える

### ■ アンプの入力光のバラツキ

- ◆ アンプの AGC (Automatic Gain Control) が有効な場合、全ての波長の増幅を効率的に行うためには、波長毎の入力光レベルをそろえる必要がある
- ◆ ぶっちゃけ同じOpticsだと、調整がなくて楽

## Optical Loss Budget を計算

区間	Optics		MUX/DEMUX		アッテネータ		プリアンプ		ダークファイバ			MUX/DEMUX		Optics				
	出力(dBm)	出力(mW)	入力	減衰	出力(dBm)	出力(mW)	入力(dBm)	増幅値	出力(dBm)	入力(dBm)	減衰	出力(dBm)	入力(dBm)	減衰	出力(dBm)	入力(dBm)		
最大許容値	dBm		-		-		-	12.00		-		-		-		-5.00		
最小許容値	dBm		-		-		-23.00			-		-		-		-27.00		
出力最大値	dBm							17.00										
ゲイン&ロス	dB			-3.00			-10.00		17.00			-12.00			-3.00			
合波					5.80	3.80	-10.00	-4.20	17.00	12.80	12.80	-12.00	0.80	0.80				
50 ch	10GBASE-LR	0.00	1.00	0.00	-3.00	-3.00	0.50	-10.00	-13.00	17.00	4.00	4.00	-12.00	-8.00	-8.00	-3.00	-11.00	-11.00
51 ch	10GBASE-LR	0.00	1.00	0.00	-3.00	-3.00	0.50	-10.00	-13.00	17.00	4.00	4.00	-12.00	-8.00	-8.00	-3.00	-11.00	-11.00
52 ch	10GBASE-LR	0.00	1.00	0.00	-3.00	-3.00	0.50	-10.00	-13.00	17.00	4.00	4.00	-12.00	-8.00	-8.00	-3.00	-11.00	-11.00
53 ch	10GBASE-LR	0.00	1.00	0.00	-3.00	-3.00	0.50	-10.00	-13.00	17.00	4.00	4.00	-12.00	-8.00	-8.00	-3.00	-11.00	-11.00
54 ch	10GBASE-LR	0.00	1.00	0.00	-3.00	-3.00	0.50	-10.00	-13.00	17.00	4.00	4.00	-12.00	-8.00	-8.00	-3.00	-11.00	-11.00
55 ch	10GBASE-LR	0.00	1.00	0.00	-3.00	-3.00	0.50	-10.00	-13.00	17.00	4.00	4.00	-12.00	-8.00	-8.00	-3.00	-11.00	-11.00
56 ch	1000BASE-LX	-1.00	0.10	-1.00	-3.00	-4.00	0.40	-10.00	-14.00	17.00	3.00	3.00	-12.00	-9.00	-9.00	-3.00	-12.00	-12.00
57 ch	1000BASE-LX	-1.00	0.10	-1.00	-3.00	-4.00	0.40	-10.00	-14.00	17.00	3.00	3.00	-12.00	-9.00	-9.00	-3.00	-12.00	-12.00

※ 「Optical Networking では常識ですYO」 (某氏)

凡例	入力パラメータ	必須
	入力パラメータ	任意
	危険パラメータ	
	チェックポイント	

TIPS : Hot Stage を行う事もトラブルの解決のためには有効



# 光アンプ + MUX/DEMUX





10Gbps × 6波長

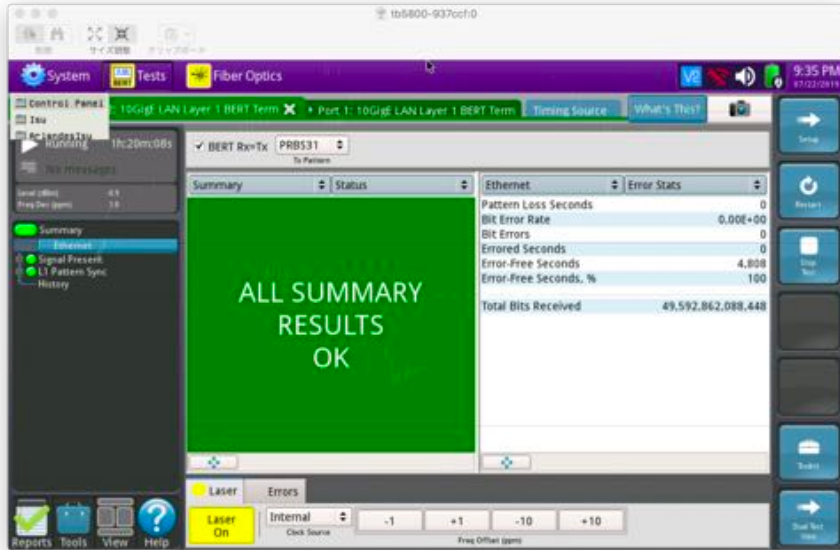
1 Gbps × 2波長

→

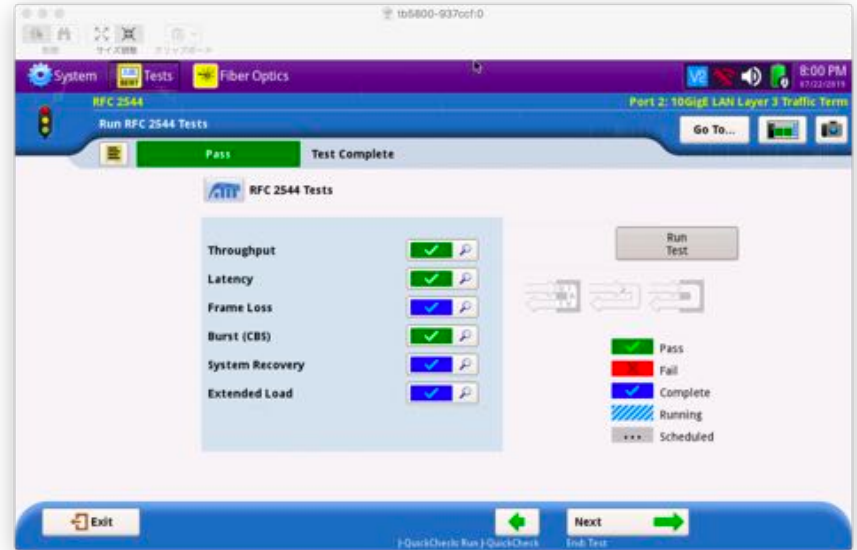
スネーク試験

→

クラウド接続



BERT試験



RFC2544



約10マイルの長距離伝送に成功

## データセンター間接続（DCI）を一から作り直す話

- 動機&ゴール：           なぜ作り直そうと思ったのか？
- 第1ステップ：           ラボ内試験
- 第2ステップ：           データセンター内接続
- 第3ステップ：           データセンター間接続
- **ネクストステップ：**   **100Gbpsオーバーの世界**

次は本格的な冒険にでかけよう

# 100Gbps以上の世界

森川さんへ

## DCIのまとめ

- ◆ DCIに用いられる技術も標準化されており、JANOGerの手の届く所にある。
- ◆ DCIも、トポロジーや距離などを割り切ることで、扱いやすくコストを下げられる。
- ◆ DCIも、ネットワークのデザインに含めるべき。

(信頼性を確保するためには、別のレイヤでのプロテクションを考えるのも重要)

**「難しい」と思っても、挑戦してみると、意外と簡単なこともある**



## community「Open Transceiver」をやっています！

### ■ 趣旨：

- ◆ インフラの100Gbps時代を迎えて、トランシーバはルータやスイッチに比較して同じくらいの投資を必要とする重要なデバイスとなっています。  
「Open Transceiver」は、トランシーバ、および、光伝送装置の運用に関してオープンに語りあう場です。

### ■ 掟：チャタムハウスルール

- ◆ セキュリティ業界でも使われているオープンな発言を即すルール
- ◆ 「個人での参加」=発言は組織を背負わない

### ■ コミュニケーションチャネル

- ◆ slack（[transceiver.slack.com](https://transceiver.slack.com)）を募集 → JANOG slack にアナウンスします

### ■ 過去のミーティング

- ◆ 2019年2月25日（月） Open Transceiver 01
- ◆ 2019年7月25日（木） Open Transceiver 02

# Acknowledgments



通称ファイバーレディ佐伯さんの  
助言がなければ、  
本プログラムは成立していませんでした。  
ありがとうございました。