

# 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」

JANOG50  
2022/7/14(Thu)

インターネットマルチフィード株式会社

菅原 淳  
sugahara@mfeed.ad.jp

## • 本プレゼンテーションについて

- アジアのIXとして初めて実施した「**マルチベンダ環境によるIX 400G相互接続実験**」についてお話しします
- 実験を通して分かった400Gの運用時に考慮すべき点やその技術的背景などについて紹介し、AS同士の対外接続における400G使用の課題などについて議論できればと考えています

## • 目次

### • インターネットマルチフィード 菅原

- IXにおける400G I/F提供の背景
- 「IX 400G相互接続実験」概要と結果について

### • インターネットイニシアティブ 竹崎さん

- IIXにおける400G導入のモチベーション
- FEC, CRCエラーカウンタについて

### • NTTコミュニケーションズ 高須さん

- NTTComにおける400G導入のモチベーション
- Degraded SERについて

### • ジュニパーネットワークス 鈴木さん

- 400G Opticsのトレンド、トラブルシューティングの勘所
- Beyond 400Gに向けて

### • アリスタネットワークスジャパン 土屋さん

- トランシーバーの電気信号のチューニングパラメータ
- インターフェースのリンクアップ時間について

### • シスコシステムズ 児玉さん

- 400G時代の新しい潮流と今後の展望

- **菅原 淳 / Jun SUGAHARA**

- インターネットマルチフィード株式会社 技術部

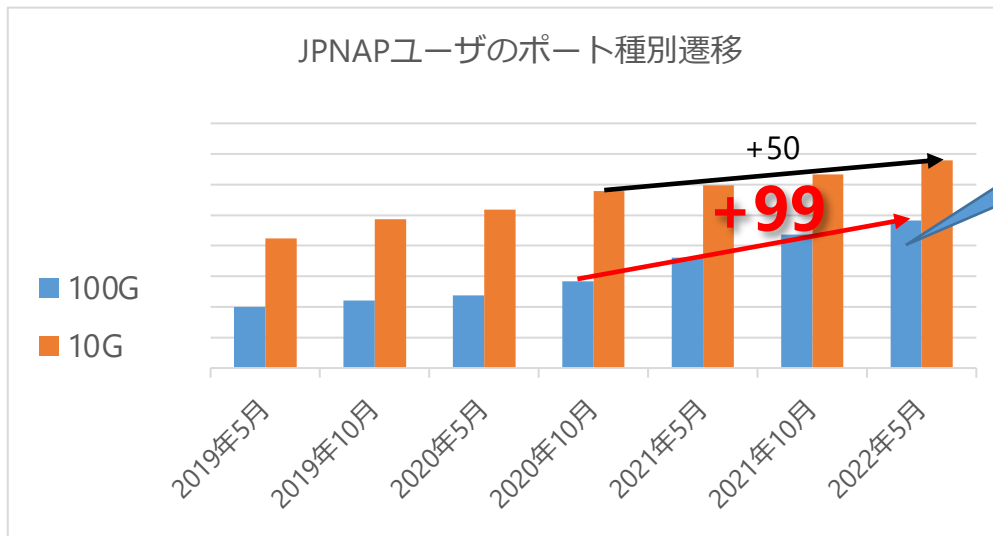
- 経歴

- 2014年
  - IIIに入社
- 2014年-2015年
  - バックボーン of 構築・保守・運用
- 2015年-2019年
  - 大阪に異動
  - 引き続き構築保守運用を続けながら運用のDR化推進など
- 2019年-
  - インターネットマルチフィードへ出向
  - JPNAPでネットワークの設計から運用などを担当
- 2022年
  - JANOG初登壇!

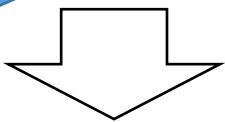


# IXから見た提供インターフェース帯域のトレンド

- インターネットエクスチェンジ(IX)とは
  - 「様々な組織のAS同士を相互接続」し効率的なトラフィック交換を実現するためのネットワーク
- ユーザに対して主に1G~100GのI/Fを提供しユーザ間のL2の接続性を提供
  - 近年のインターネットトラフィック増により100G I/Fの提供数は10Gを上回る勢いで増加中



さらに100Gユーザのうち  
1/3が複数本LAGでの利用



国内でもより広帯域なI/F  
の需要が高まっている

→ISP視点での広帯域なI/F導入のモチベーションなどは竹崎さん、高須さんからお話しいただきます

- 100Gより広帯域なI/Fとして2017年に400G-LR8,FR8が規格化
  - ただし価格は100Gと比較して非常に高価・・・
  - 昨年になり4波CWDMを用いたより安価な規格が登場

Standard	Name	Maximum Distance	Modulation	Wavelength	Fiber	Approved
IEEE802.3bs	400GBASE-LR8	10km	PAM4	LANWDM	SMF	2017
IEEE802.3bs	400GBASE-FR8	2km	PAM4	LANWDM	SMF	2017
100G Lambda MSA	400GBASE-LR4-10	10km	PAM4	CWDM	SMF	2021
IEEE802.3cu	400GBASE-FR4	2km	PAM4	CWDM	SMF	2021

→その他の400G規格や展望については鈴木さん、児玉さんよりお話しいただきます

- **海外IXにおける400G I/Fの提供状況**

- 2019/03 DE-CIX(ドイツ)世界で初めて400G I/F提供開始
- 2019/11 Netnod(北欧) 複数ロケーションで400G I/F提供開始
- 2021/09 France-IX(フランス) で400G I/F提供開始

- **アジアや日本国内では**

- アジア圏では400G I/Fを提供しているIXは確認できず
- 400G-FR4,LR4などのより安価な規格の登場やユーザの現況なども踏まえてJP NAPでも本格的な検討を開始

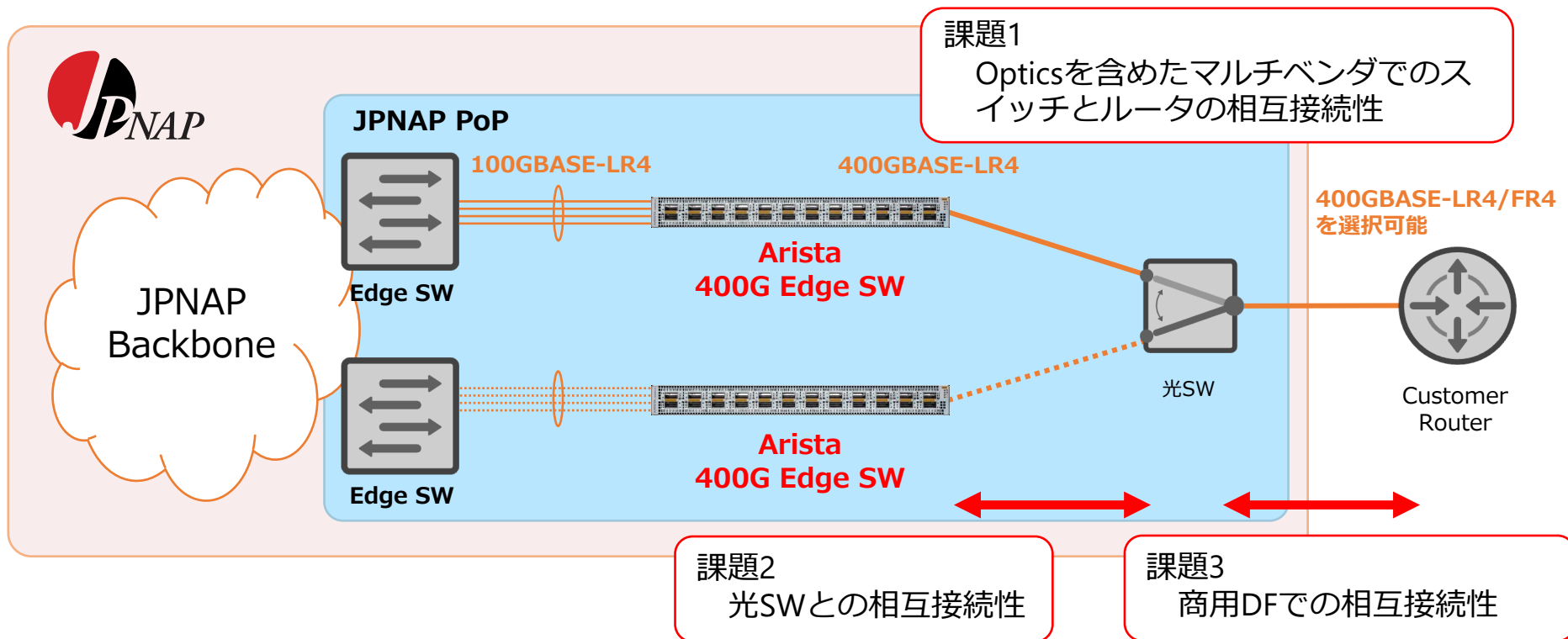
今すぐ提供準備を始めよう！

# IXにおける400G I/F提供の課題



## 提供構成のポイント

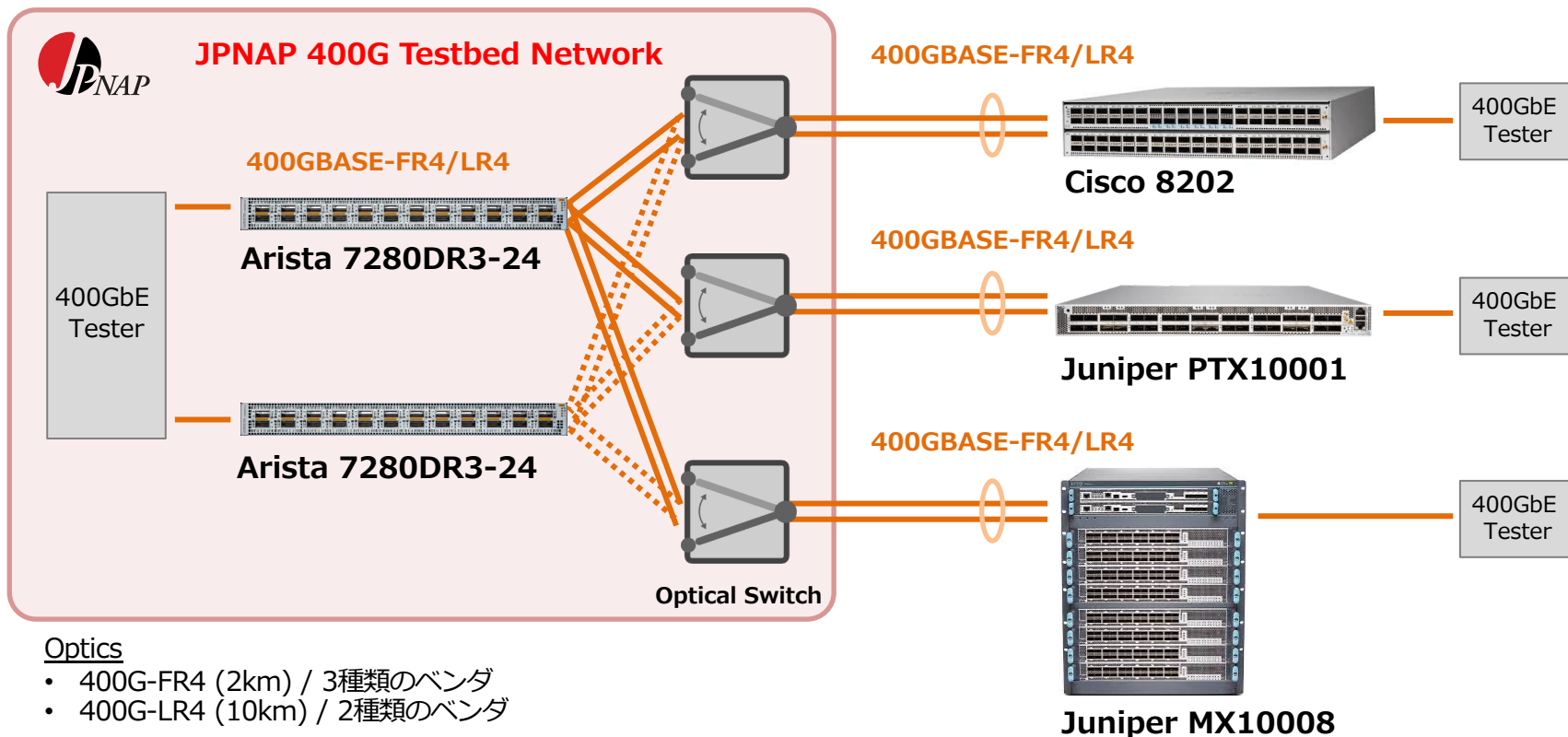
- 既存の100GエッジSWから400GエッジSWを張り出し
- 従来の品質を維持するため光SWによる冗長化
- IXスイッチ側は400G-LR4に統一しユーザ側でLR4/FR4を選択可能に



※構成概略図です

# ラボにおける相互接続実験

- MFのラボ環境に各社機器を持ち寄り実験を実施
  - A) マルチベンダ環境における物理的なLinkUp確認
  - B) 光SWの切替え動作の確認
  - C) BGP確立と試験機を用いた400Gトラフィック導通の確認



## ・マルチベンダによる物理LinkUp試験

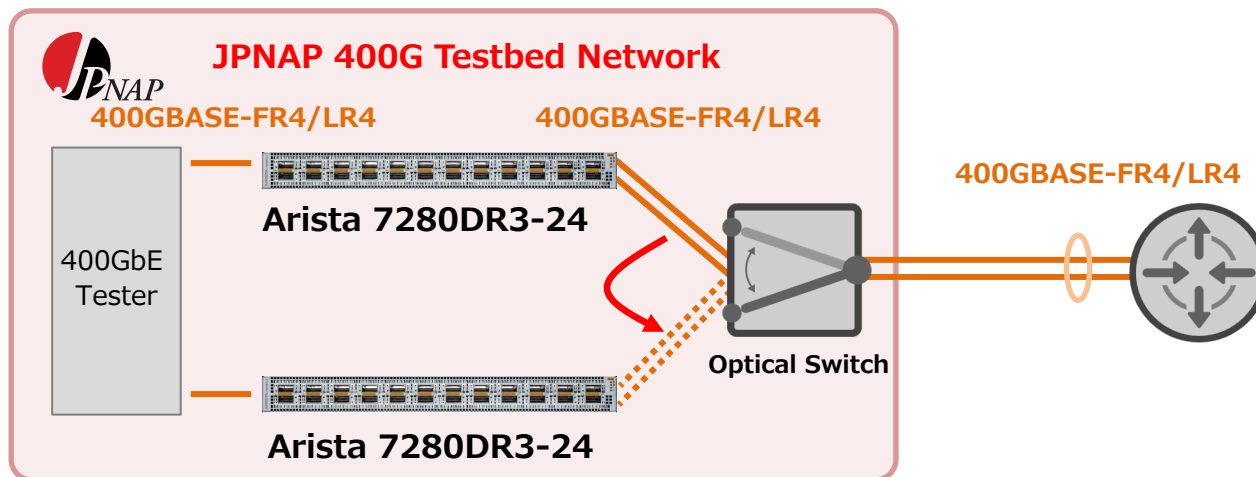
✓異なるベンダのスイッチ・ルータ・Optics間でLinkUpを確認

マルチベンダ同士でのLinkUp試験結果				
		Cisco8202 Juniper PTX/MX		Cisco8202
		400G-LR4 (ベンダA)	400G-FR4 (ベンダA)	400G-FR4 (ベンダC)
Arista	400G-LR4 (ベンダA)	✓	✓	✓
	400G-FR4 (ベンダA)	✓	✓	未実施
	400G-LR4 (ベンダB)	✓	✓	未実施
	400G-FR4 (ベンダB)	✓	✓	未実施

✓ 異ベンダ400G-LR4↔400G-FR4同士でも問題なくLinkUp  
IX側は収容ポートのLR4統一が可能となりユーザ側でFR4ま  
たはLR4のOpticsを選択可能

## 光スイッチによる400G I/F光の認識と自動切替試験

- ✓ 光スイッチにおいて光レベルを正常に認識し自動切替も動作
- ✓ ただし、Juniper MX10008との接続において、稀に切り替え後のLinkUpに100秒程度かかる事象を確認



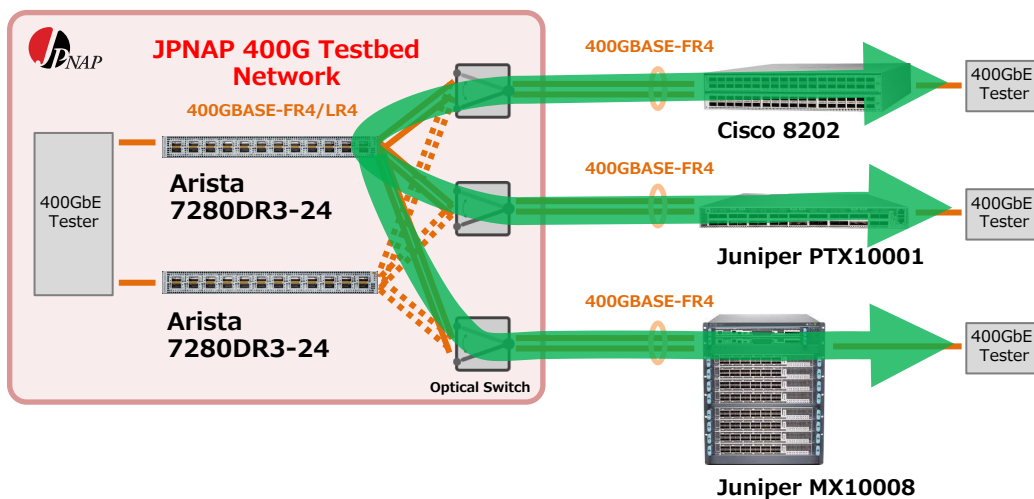
✓ 400G I/Fにおいても光SWによる筐体冗長の基本動作には問題なし

✓ 一部機器で切り替え後のLinkUpに遅延が生ずる事象を確認

→メーカーにもご協力いただいた再試験結果については、鈴木さん、土屋さんよりお話しいただきます

## 400G試験機を用いたL3トラフィック導通試験

- ✓ Cisco,Juniper間で400G LAGにおいて問題なくBGPピア確立が可能
- ✓ 400GのL3トラフィックをIX経由で問題なく導通可能
- ✓ 試験機と一部ベンダOpticsの組み合わせにおいてFEC Symbol Error Rateの上昇を確認



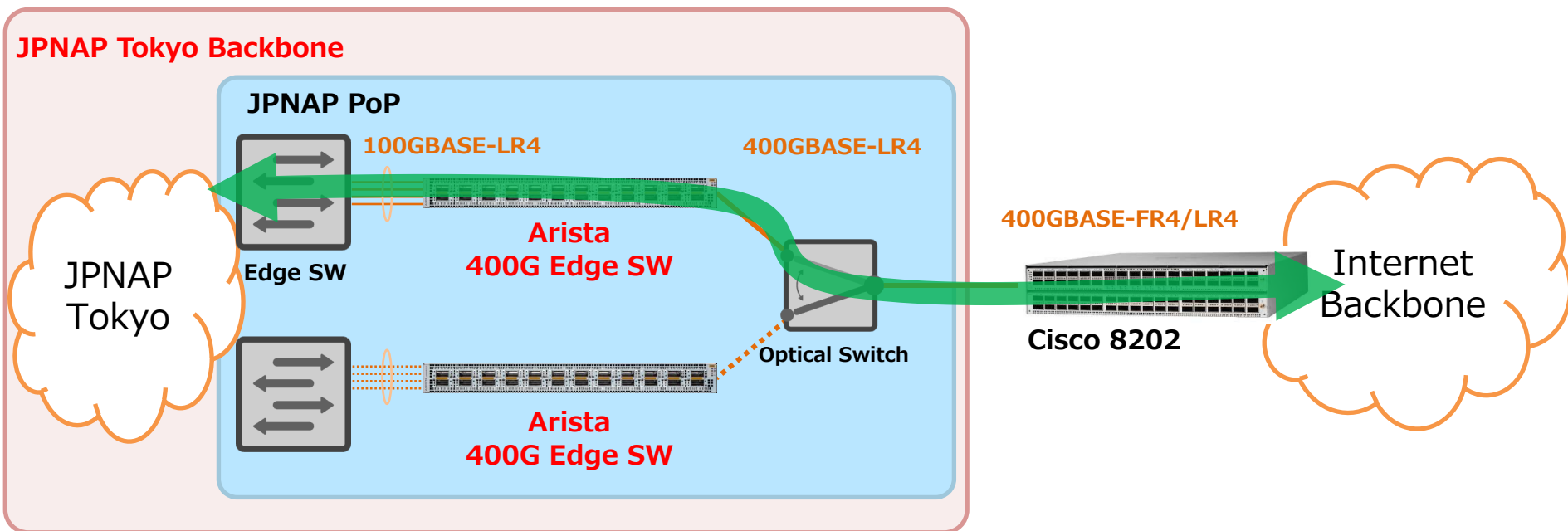
✓ LAGやBGP経路交換などは10G/100G時代と変わらず運用が可能

✓ 試験機と特定ベンダの400G Opticsの組み合わせにおいて信号品質の劣化を確認

→信号品質の監視や品質劣化の原因については、高須さん、土屋さんよりお話しいただけます

# 商用環境における実験

- **JPNAP PoPまでの商用ダークファイバを用いた相互接続実験**
  - 商用DFを用いたLinkUpと商用インターネットトラフィック導通試験





- **JPNAP PoPに敷設したダークファイバのスペック**

- 距離: 1.73km、減衰: 3.3dBm@1310nm

- **LinkUp試験結果**

DF両端でのLinkUp試験 (1.73km/3.3dBm減衰)			
	400G-LR4 (Vendor A)	400G-FR4 (Vendor A)	400G-FR4 (Vendor C)
400G-LR4 (Vendor A)	✓	✓	✓

✓ 日本の商用DF環境においてもFR4~LR4によるLinkUpが可能  
1.7kmというFR4の上限(2km)に近いレンジでも十分な性能

- **商用インターネットトラフィック導通試験**

- 400G I/Fでも品質劣化なくトラフィック導通が可能  
→結果の詳細は竹崎さんパートにて

- **400Gにおけるマルチベンダ相互接続性に重大な問題はなかった**
  - 一部試験機とOpticsの組み合わせで信号劣化が生ずる可能性があり  
事前確認の上そういった組み合わせは避ける必要がある  
→信号劣化確認や劣化原因は高須さん、土屋さんパートにて
- **光SWは従来通り利用可能**
  - 一部切替後のLinkUpに時間がかかるという事象が確認されたが、追加実験の結果光SW起因の可能性は低
  - やはりOpticsの組み合わせには注意  
→詳細は鈴木さん、土屋さんパートにて
- **商用DFを用いた相互接続性も問題なし**
  - 実インターネットトラフィックの導通も問題なし  
→実験の詳細は竹崎さんパートにて

**これら実験を通して400G I/Fの商用提供は可能と判断**

# IX視点から見た400Gで変わること

- **400GはLinkUpにかかる時間が長くなりました**
  - メーカーに関わらずおよそ100Gの約4倍の8秒～12秒！
  - LinkFlapの発生によりBGPもDownすると復旧するまでには早くても約10秒は必要
  - しかも影響するトラフィックは100G時代の最大4倍
- **光SW切替などによる瞬間的な光断によるLinkDown回避はHold Down Timerを調整することで可能**
  - Juniper, Ciscoにおいては**12000msec**でLinkDown回避可能

推奨されるHold Down Timer値	
1G～100G	100msec～2000msec
<b>400G</b>	<b>12000msec</b>

本当に設定すべき・・・？

是非ご意見ください

**続いてIII 竹崎さんよろしくお願ひいたします！**

- **400GのHold Down Timer 12000msecは設定すべき？**
- **伝送品質の監視について**
  - **Telemetry? SNMP ?**
- **Degraded SERの導入要否について**