
超高速・低消費電力光ネットワーク技術の研究開発 (課題 I (c) プロトコル無依存リンク多重化技術)

2015.10.7

株式会社日立製作所 研究開発グループ

研究代表者 坂本 健一

研究分担者 豊田 英弘

課題 I (c) プロトコル無依存リンク多重化技術

Contents

1. 研究開発内容
2. 研究開発成果
3. 成果展開と波及創出の取り組み
4. まとめ

課題 I (c) プロトコル無依存リンク多重化技術

Contents

1. 研究開発内容
2. 研究開発成果
3. 成果展開と波及創出の取り組み
4. まとめ

1.1 アクセスネットワークの課題と本研究の狙い

■ アクセス(加入者・局舎)NWの課題

■ データセンタ内/間ネットワーク

- 多種多様なリンク層プロトコルの利用し、物理的に分離されたLAN/SANで構成

■ L2/L3スイッチによる多段のネットワーク構成

- フレーム多重時の転送遅延が増加し、ネットワーク全体の処理パフォーマンスを悪化

- 400Gbps以降、フレーム多重のワイヤレート処理が困難となり、実現コスト・電力コストが急増

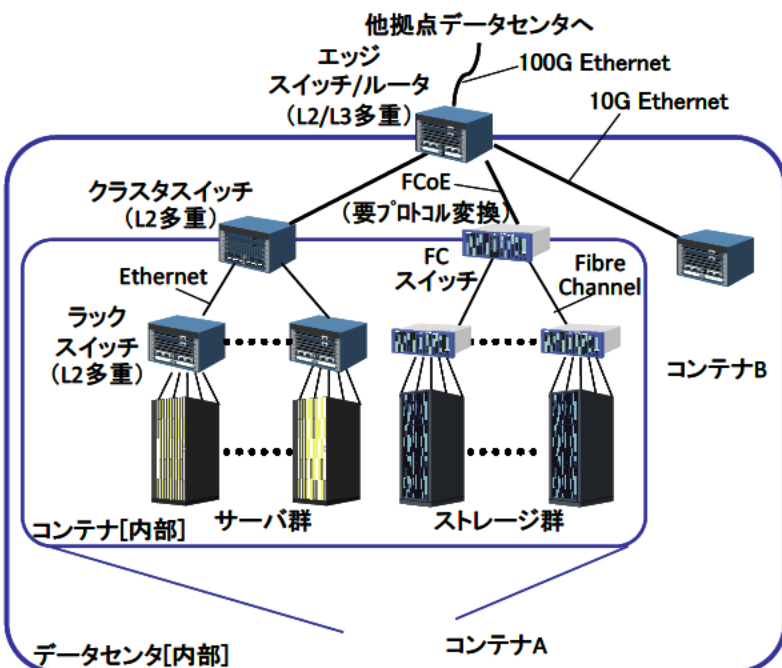
■ 本研究の狙い

- リンク層プロトコルや伝送レートに依存しない、物理リンク多重化方式(L1多重)を提案

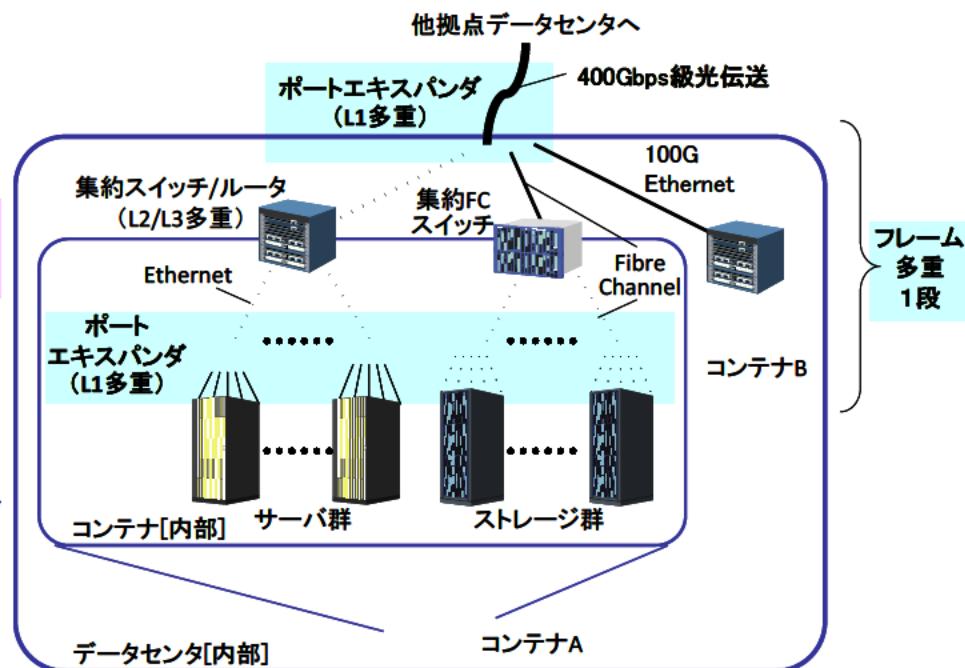
- 末端收容にL1多重を適用、L2/L3処理を集約し、不必要なフレーム多重処理を削減して低消費化

- フレーム多重段数の削減、多重処理の低遅延化により、センタ全体の処理パフォーマンスを改善

- LAN/SANを物理層レベルで統合可能とし、ネットワーク装置台数を削減



従来手法によるデータセンタ構成



本提案によるデータセンタ

I アクセスネットワーク高速大容量化・低消費電力化技術

(b) 光多値伝送向け高性能信号処理技術（日立、NTT）

(c) プロトコル無依存リンク多重化技術（日立）

ア) 物理リンク多重化技術

上位層プロトコルに依らず、性質の異なる複数の物理層リンクを、物理層が扱うビットデータの状態で多重化する技術の確立

イ) 伝送レート透過・再生技術

様々な伝送レートや、周波数誤差を持つ非同期レートのリンクを多重・分離可能とする、多重一分離装置間での正確な伝送レート同期・再生技術の確立

ウ) 物理リンク管理技術

宛先情報を持たない物理層リンクを、多重装置と分離装置間で効率良く管理する技術の確立

(d) マルチコアファイバ光接続技術（日立）

1.3 プロトコル無依存リンク多重化技術

■ 伝送するリンクのプロトコルや伝送速度に依らず、かつ低消費電力、高効率、低処理遅延な多重化方式

■ サブ課題 ア) 物理層リンク多重化技術

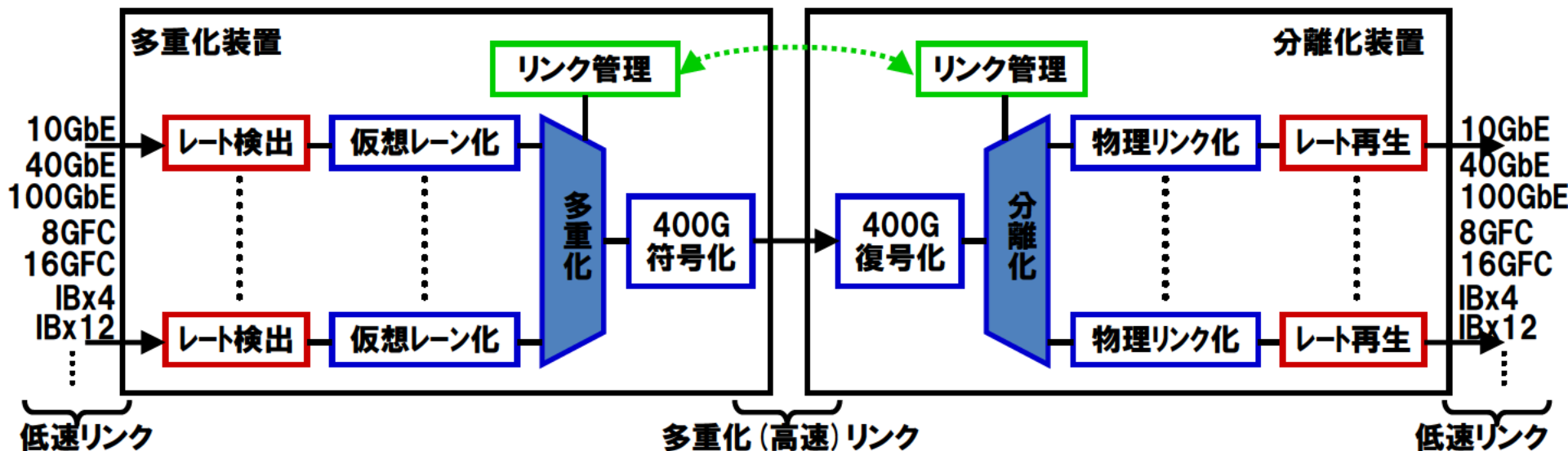
■ マルチレートの低速リンクを多重化し、固定レート的高速リンクへ変換

■ サブ課題 イ) 伝送レート透過・再生技術

■ 多重装置(受信部)ー分離装置(送信部)間で、低速リンクの伝送レートを正確に同期・再生

■ サブ課題 ウ) 物理層リンク管理技術

■ 宛先情報を持たない物理層リンクを、多重装置と分離装置間で効率良く管理



■ プロトコル無依存、かつ伝送レート無依存であり、またこれら多数の低速リンクを論理多重し、400Gbpsの大容量リンクに重畳する多重化技術を確立

■ 2種以上の異なるリンク層プロトコルに対応

■ 2種以上の異なる伝送レートに対応

■ 異なる源信(非同期源信)から生成されたリンクに対応

■ Ethernet: 10GbE(10.3125Gbps)、40GbE(10.3125G×4)、
100GbE(10.3125Gbps×10、25.78125Gbps×4)

■ FibreChannel: 4GFC(4.25Gbps)、8GFC(8.5Gbps)、10GFC(10.51875Gbps)、16GFC
(14.025Gbps)

■ InfiniBand: DDR(5.0Gbps)、QDR(10.0Gbps)、FDR(14.025Gbps)、EDR(25.78125Gbps)
...等

■ 低速リンク 8本以上を多重可能(多重化本数は伝送レートに依存)

■ 例: 低速リンク 40Gbps×8本(以上) ⇔ 高速リンク 400Gbps

■ 多重化処理遅延: 1マイクロ秒以下(数100ナノ秒オーダーを目標)

■ 本方式による試作機を開発し、多重伝送実験により提案方式の有効性を実証

■ 他課題とも連携し、統合実証実験を実施

■ 課題 I (b)「光多値向け高性能信号処理技術」: 高速リンクに400Gbps光I/Fを利用

■ 課題 I (d)「マルチコアファイバ光接続技術」: リンク接続にマルチコアファイバを利用

課題 I (c) プロトコル無依存リンク多重化技術

Contents

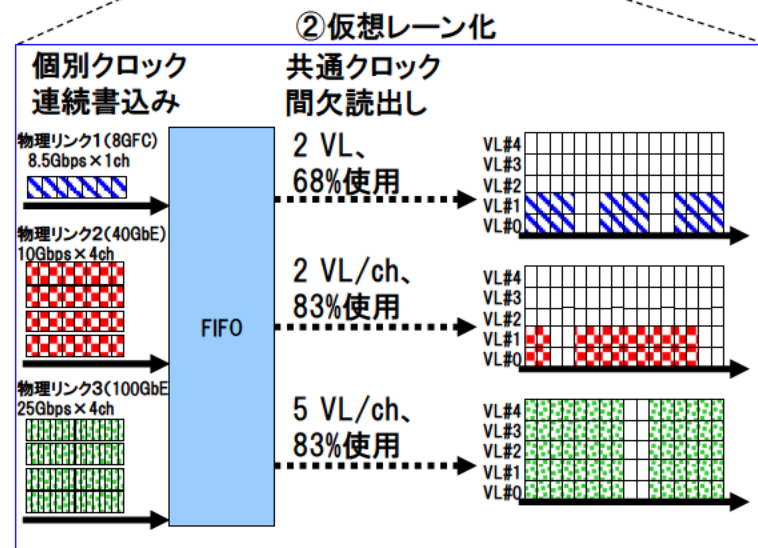
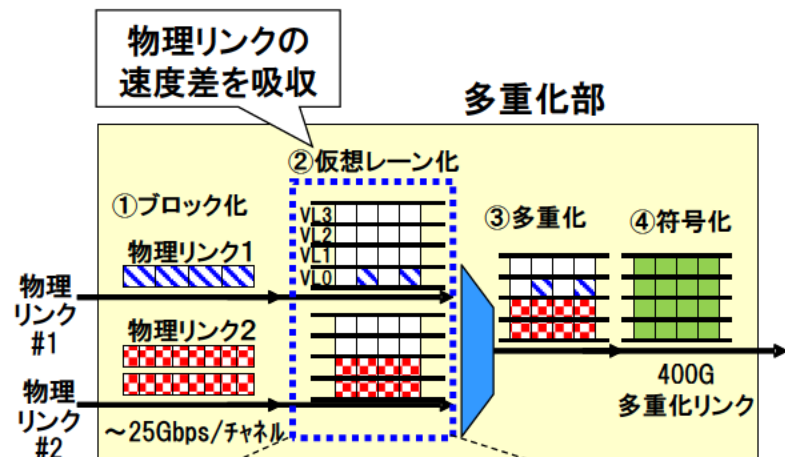
1. 研究開発内容
- 2. 研究開発成果**
3. 成果展開と波及創出の取り組み
4. まとめ

2.1 物理リンクの収容と多重化方式

- サブ課題 ア) 物理層リンク多重化
 - マルチレートの低速リンクを多重化し、固定レート的高速リンクへ変換
 - 「プロトコル無依存」→データフォーマット、伝送レートを維持したトランスペアレント伝送

■ 多重化方式

- ① ブロック化
 - 受信したビット列データを固定長(64bit)に分割
- ② 仮想レーン化
 - 低速リンクから仮想レーン(=6.25Gbps)への変換
伝送レート < 6.25Gbps × n (n ≤ 5)
- ③ 多重化
 - 仮想レーン(全64本=400Gbps)をビット多重
未使用のレーンは多重化に含めない
→使用レーン数を分離化側へ通知
- ④ 符号化
 - 64B66B符号、CAUIベースのデスクュー方法
 - 400Gbps一括スクランブル



物理リンク多重化部の動作

2.2 分離処理と非同期リンクのレート再生

■ サブ課題 イ) 伝送レート透過・再生

- 多重装置(受信部)-分離装置(送信部)間で、低速リンクの伝送レートを正確に同期・再生
- 非同期源信のリンク多重化を実現

■ 分離化、伝送レート再生

① 復号化

- 多重化の符号化の逆変換

② 分離化

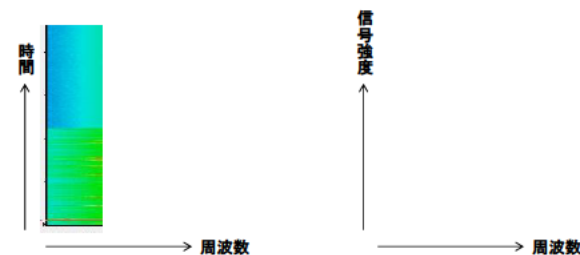
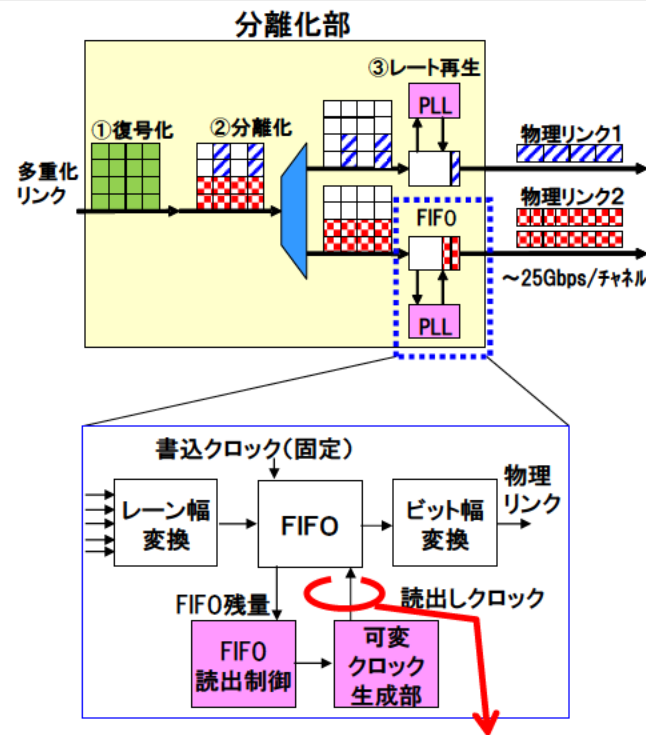
- リンク管理情報(後述)を元に、仮想レーンから物理リンクを分離

③ レート再生

- 各仮想レーンの有効ブロックをFIFOに格納(共通の書き込みクロック)
- FIFO残量を元に、読出しクロックの周波数を調整
- FIFOを連続読み出し(各チャネルの再生クロック)

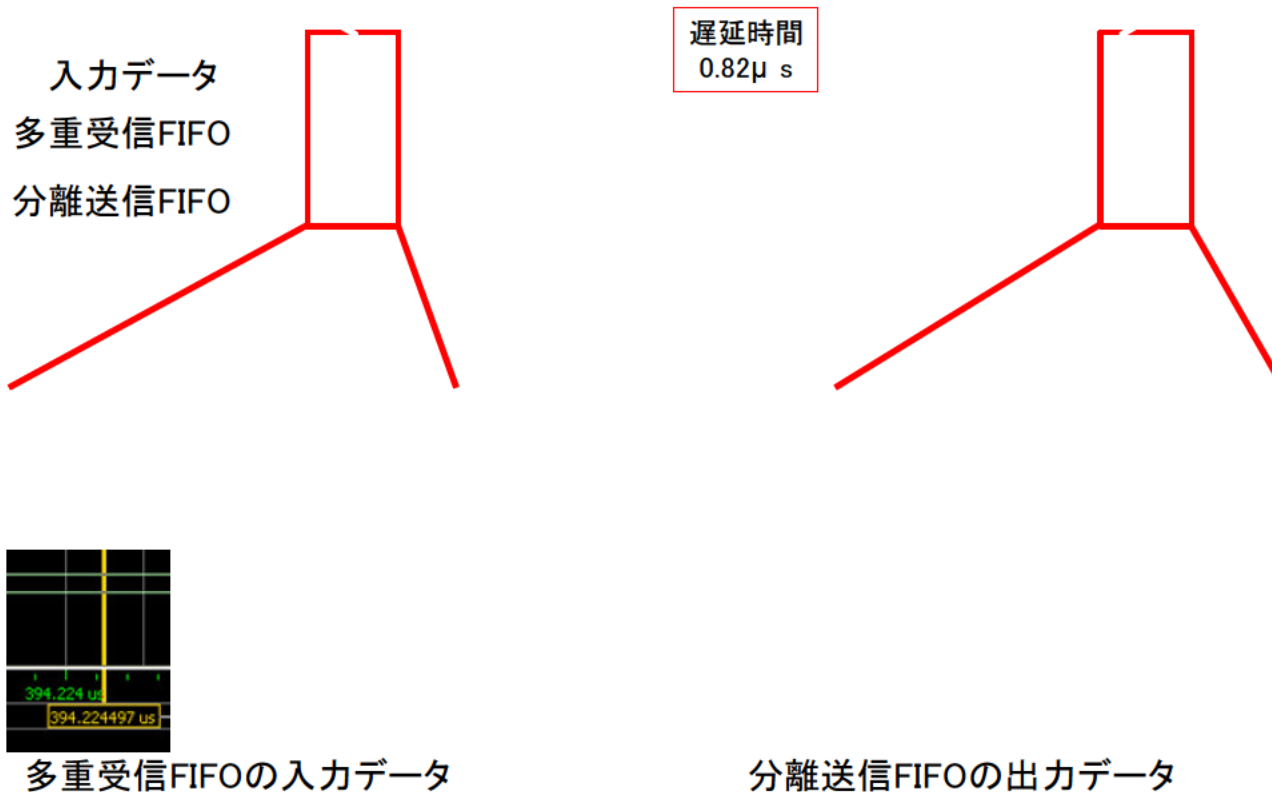
■ 再生クロック

- VCOの入力電圧をDACで制御し、出力を平滑化
→ユーザリンクの受信CDRが追従できる変化速度
- $\pm 400\text{ppm}$ の非同期リンクへ追従



2.3 多重化処理遅延時間の検討

- 多重化～分離化の処理遅延目標: $< 1\mu\text{s}$
- シミュレーションにて処理遅延時間を確認
 - 処理遅延 $0.82\mu\text{s}$ (搭載FPGA間の転送、通信路の遅延を含まず)



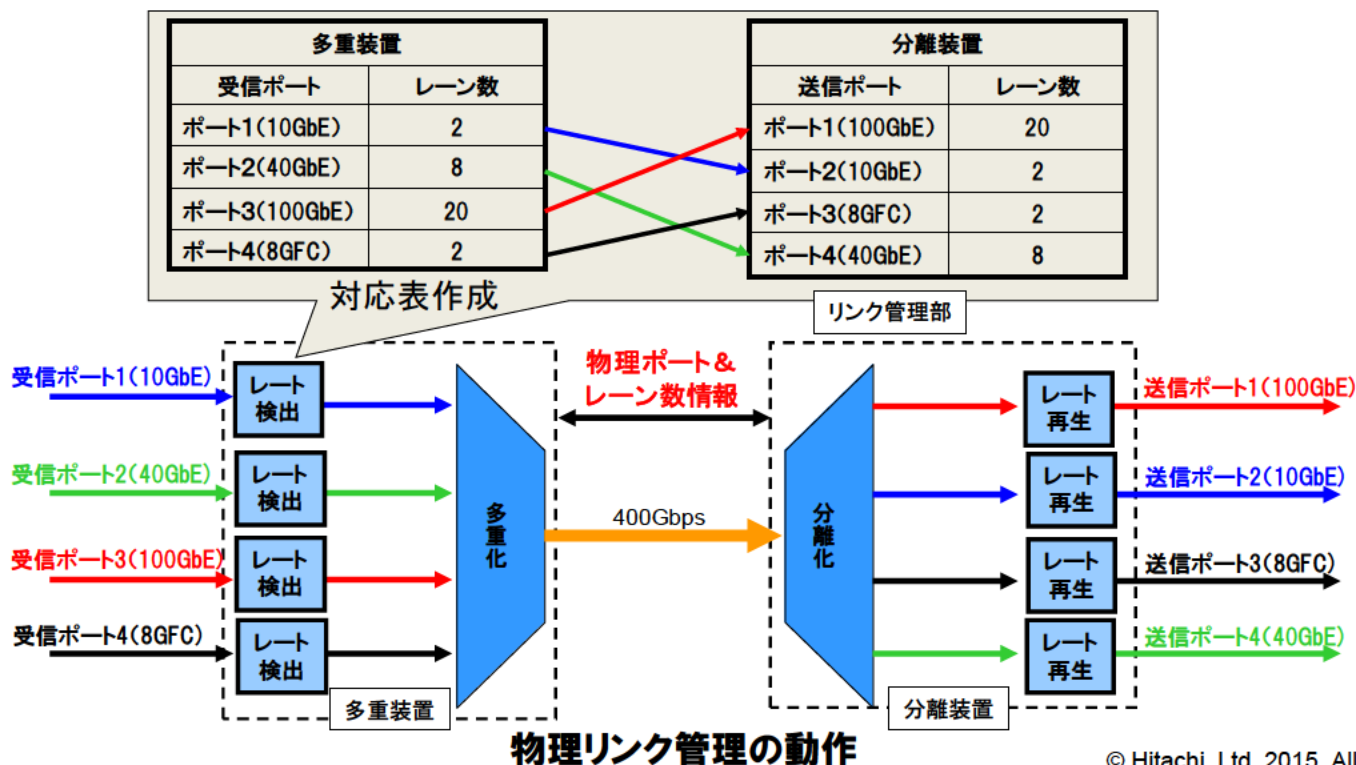
2.4 物理層リンク管理の実装

■ サブ課題 ウ) 物理層リンク管理技術

- 目宛先情報を持たない物理層リンクを、多重装置と分離装置間で効率良く管理

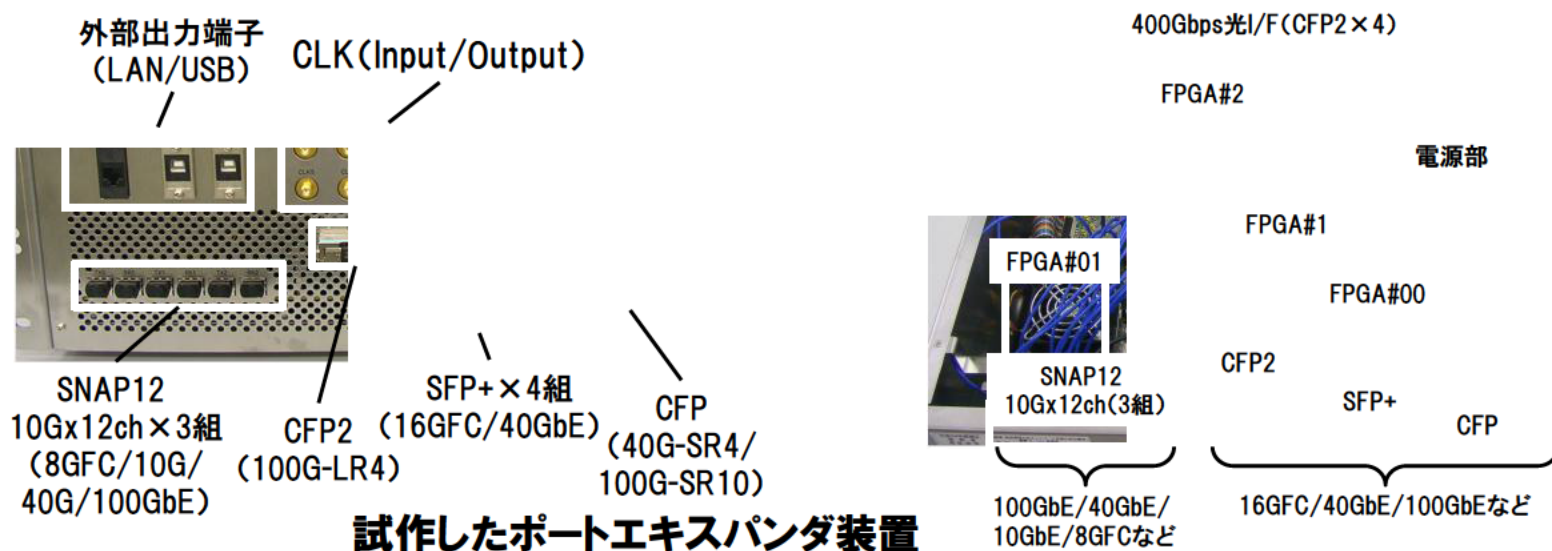
■ ポートマッピング

- リンク管理を対向装置間でポート対応付けをネゴシエーション
 - 物理リンク-仮想レーン間マッピング
 - 多重側-分離側(及び全二重リンク)間マッピング
- ポート間の入れ替え、ポート増減に柔軟に対応



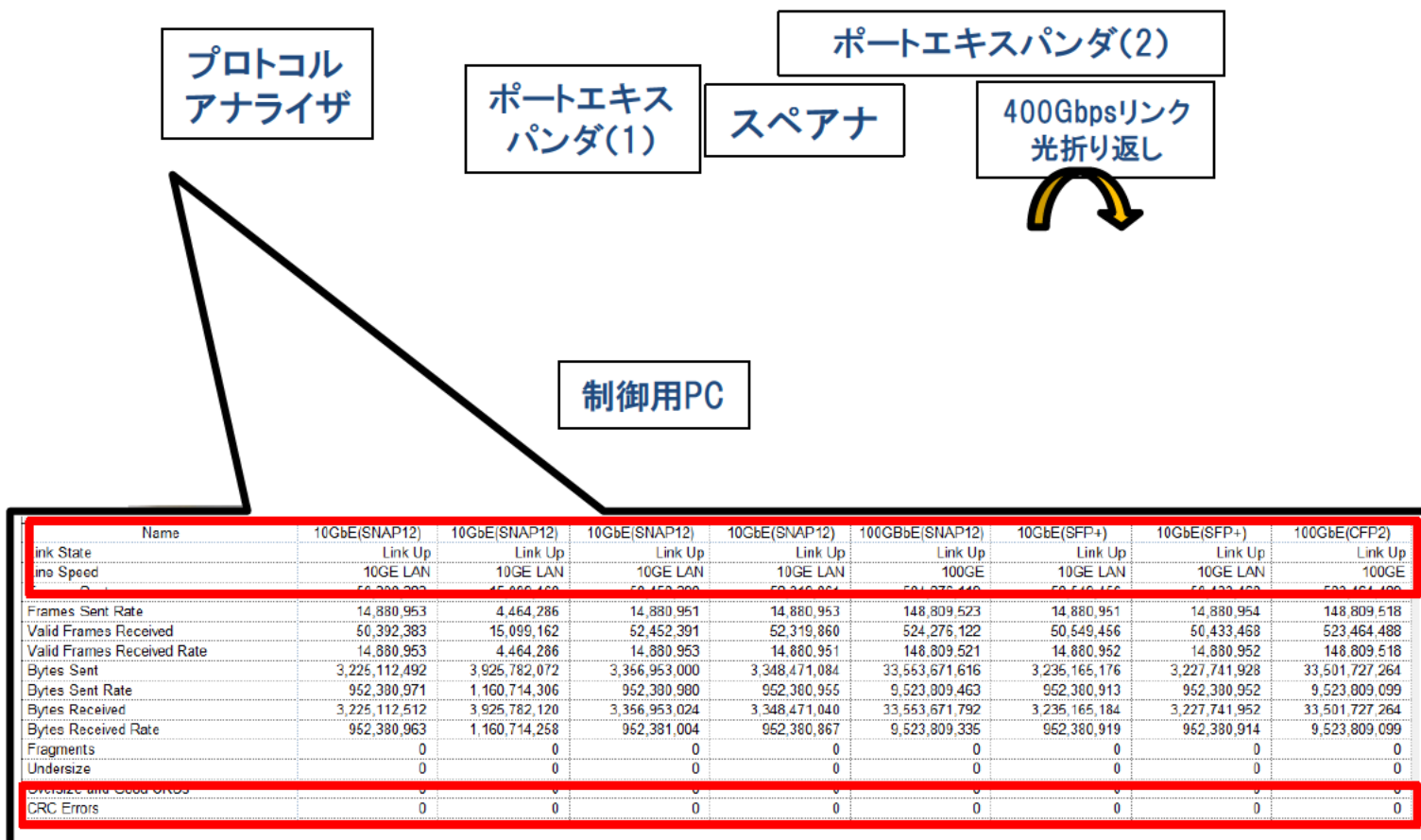
2.5 試作機開発

- 搭載光モジュール: CFP2、CFP、SFP+ ×4組、SNAP12(10G×12ch)×3組
- 対応IF: FibreChannel(8G、16G)、Ethernet(10G、40G、100G)、Infiniband
- 使用FPGA: 計4個
 - FPGA#00(XC7VH1140T): 25G系物理リンクの収容
 - FPGA#01(XC7VH1140T): 10Gbps以下の物理リンクの収容
 - FPGA#1(XC7VH1140T): 多重化論理
 - FPGA#2(XC7VX870T): 400G多重化I/F(25Gbps×16ch、CFP2×4台と接続)
 - FPGA実装率: #00(82.3%), #01(53.9%), #1(65.9%), #2(84.8%)
- 消費電力: 約350W(FPGAの消費電力は117W)



2.6 予備実験による方式検証

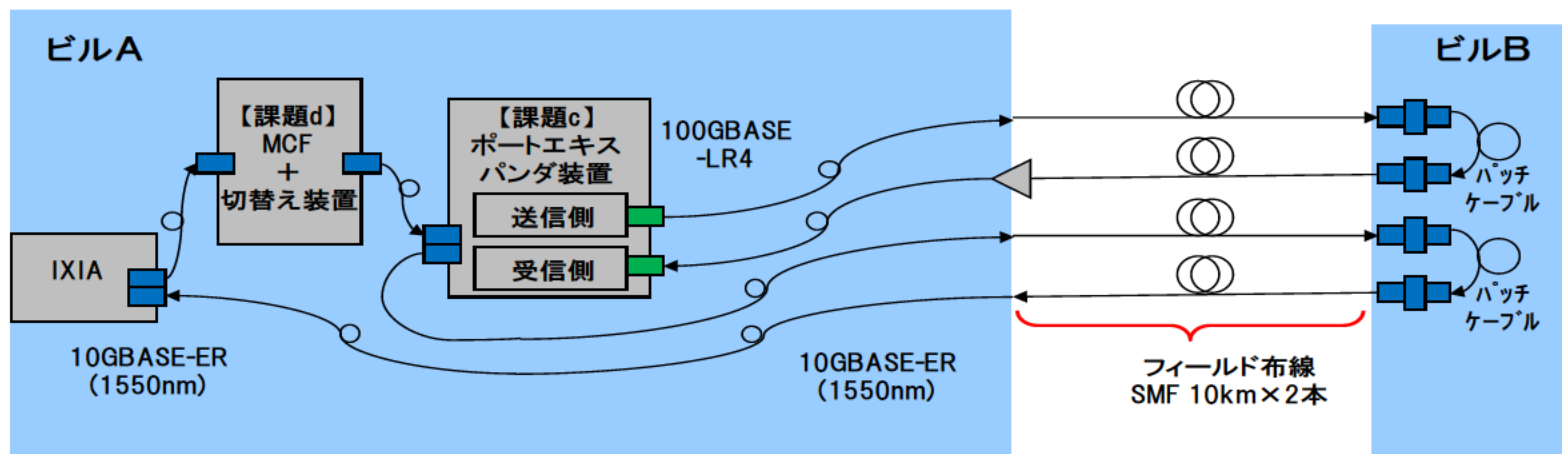
- 8ポート以上のユーザポート(6×10GbE, 2×100GbE)から400Gbpsへの多重化を検証



8ポート多重化の実験

■ 連携実証実験(H27/2/23-3/6実施):

- フィールド敷設の光ファイバ(20km)を用いた伝送実験を他課題(I -b、I -d)と実施
- 多重分離の長距離伝送を実証
 - 多重化リンクの送受信間ネゴシエーションに対する20km長距離伝送の影響
 - 伝送経路が災害等により断線した際のリンク瞬断発生時の影響



フィールド実証実験(課題I(d)連携)

課題 I (c) プロトコル無依存リンク多重化技術

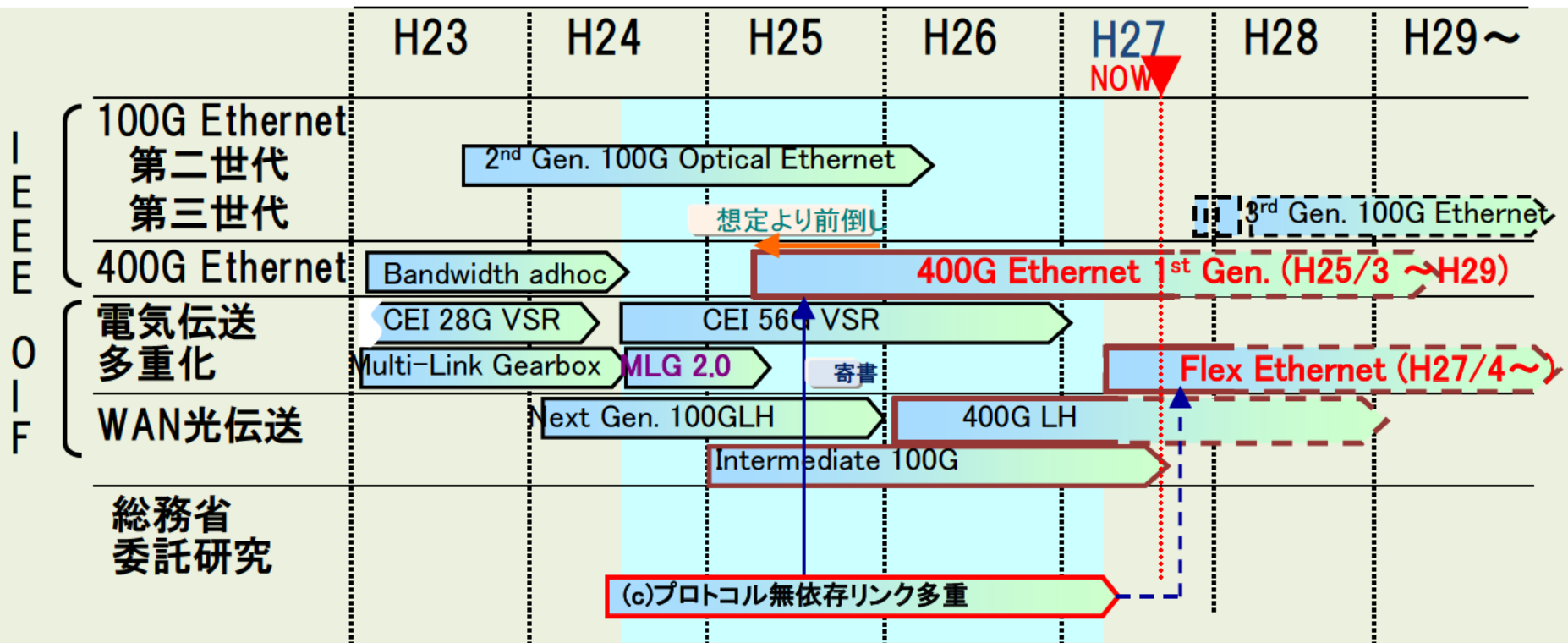
Contents

1. 研究開発内容
2. 研究開発成果
- 3. 成果展開と波及創出の取り組み**
4. まとめ

3. 1 標準化の位置付け

- 本プロジェクトの研究成果(符号化技術、多重化技術)を、IEEE、OIF等へ展開

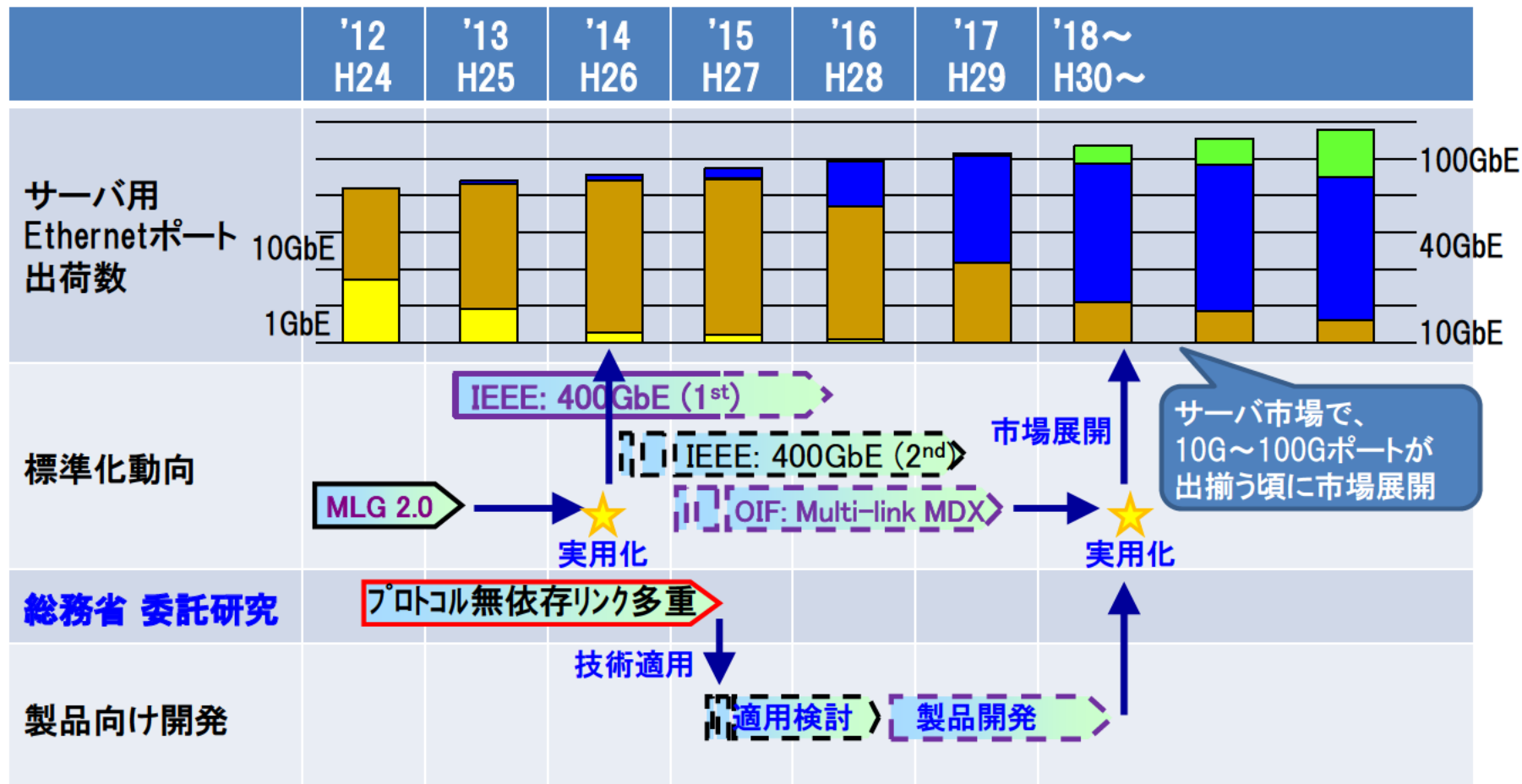
標準化委員会	標準化対象
IEEE802.3	光インタフェース、通信プロトコル(L1,L2)
OIF MLG, FlexEther	光/電気インタフェース、多重化方式(L1)



OIF: optical internetworking forum
CEI: common electrical interface

3. 2 市場動向と実用化時期

- 400G世代の標準化完了後、サーバ市場で100G世代ポートが普及する頃に市場へ展開



- 様々なリンク層プロトコルを低遅延に一括して多重化し、400Gbpsの高速リンクを効率的に利用可能とする、プロトコル無依存リンク多重化技術を開発
- データセンタ内/間ネットワークの処理効率、使用電力効率を向上する、新たな多重化方式を提案
- 提案方式を搭載したポートエキスパンダ装置の試作、実証実験

HITACHI
Inspire the Next