

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)

The Research and Development Project for the Ultra-high-speed and Green Photonic Networks
(Protocol-independent link-multiplexing technology)

研究代表者

坂本健一 株式会社日立製作所 研究開発グループ
Kenichi Sakamoto, Research & Development Group, Hitachi Ltd.

研究分担者

豊田英弘 Hidehiro Toyoda	光野正志 Masashi Kono	堤聡 Satoshi Tsutsumi	高瀬誠由 Masayuki Takase
	株式会社日立製作所	研究開発グループ	
	Research & Development Group, Hitachi Ltd.		

研究期間 平成 24 年度～平成 26 年度

概要

インターネットの利用拡大に伴う通信量と消費電力の急増に対処する超高速・低消費電力な光ネットワーク技術の実用化が急務となっている。本研究では、データセンタ内のネットワークにおいて、多数の低速インタフェースを束ねて、電力的・物量的に高効率な 400Gbps インタフェースに多重化する「プロトコル無依存リンク多重化技術」を開発した。通信方式の提案及び試作機を開発し、敷設ファイバを用いた実証実験を行い、提案技術の実用性を示した。

1. まえがき

今日、ICT 利活用の増進に伴う通信量及び消費電力の急激な増大に対応するため、伝送方式の高性能化や新型ファイバの導入等により、ネットワーク全体の超高速化、低消費電力化、耐災害性の強化を同時に実現する技術を確認し、国民生活の利便向上と地球温暖化対策が急務となっている。

我が国のインターネット通信量は大幅な伸びを続けており、今後も大幅な増加が予想されている。この通信量の伸びに対応するために、これまでの通信機器を単純に高速化した場合、伝送する情報量の増加に比例して通信機器の消費電力も大幅に増加することとなる。そのため、大量の情報を高速かつ低消費電力で伝送できる通信方式や通信機器が求められている。また、災害時等におけるネットワークの途絶といった通信環境の激変に対しても、必要な通信を維持できる、柔軟性の高いネットワークの構築も必要とされてきている。これらの課題を解消するため、伝送方式の高性能化や新型ファイバの導入等により、ネットワーク全体の超高速化、低消費電力化、耐災害性の強化を同時に実現する技術を平成 26 年頃までに確立し、国民生活の利便向上と地球温暖化対策に貢献することが本研究全課題の目的である。さらに、本技術の国際標準化を推進することにより、我が国の光ネットワーク技術の国際競争力向上に資する。

また更に課題 I (c) 「プロトコル無依存リンク多重化技術」では、複数の加入者ネットワークを収容し、ユーザが使用する様々なリンク層プロトコル (IEEE802.3 規格、FibreChannel、InfiniBand 等) が混在する局舎ネットワークにおいて、加入者ネットワークにおける最大 100Gbps 級の伝送を 1 本の大容量リンクに束ねることにより、400Gbps 級の高速大容量伝送を有効活用する技術の開発が目的となる。本課題では、2 種以上のリンク層プロトコル (IEEE802.3 規格、FibreChannel、InfiniBand 他)、2 種以上の通信速度 (10G/40G/100Gbps IEEE802.3 規格、8G/16Gbps FibreChannel、DDR/QDR Infiniband 等) のリ

ンクを 8 本以上論理多重し、400Gbps (マルチレーン伝送) の大容量リンクに多重化する多重化方式を立案し、多重化処理を 1 マイクロ秒程度の処理遅延で実現する。また本方式を実装した試作機を開発し、上記プロトコルの多重伝送実験を行って動作実証する。

2. 研究開発内容及び成果

プロトコル無依存リンク多重化技術を確認するため、平成 24 年度及び 25 年度には、以下 3 つのサブ課題を設定し、要素技術の研究開発を実施した (図 1)。本課題では、要素技術・方式の基礎検討、仕様決定を行い、一部機能の試作を実施した。

課題ア) 物理リンク多重化技術：

2 種以上のリンク層プロトコル、2 種以上の伝送レートのリンク 8 本以上を、400Gbps 以上の大容量リンクに多重化可能なリンク多重化方式。

課題イ) 伝送レート透過・再生技術：

5Gbps 帯から 25Gbps 帯の伝送レート (チャンネル毎)、±100ppm の周波数または伝送レート誤差を持つ非同期リンクからのデータを受信し、これを固定レートの伝送を通過した後も、同等な伝送レートの再生を可能とする伝送レート透過・再生方式。

課題ウ) 物理リンク管理技術：

低速度の物理リンクと複数本の固定レートの仮想レーンとのマッピング方式、及び受信ポート (多重装置) と送信ポート (分離装置) 間でのポートマッピング同期方式。

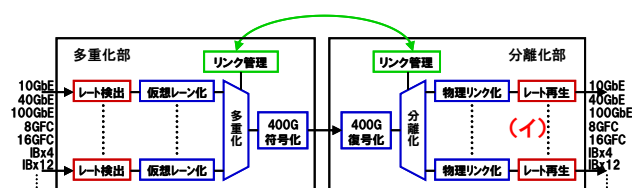


図 1： プロトコル無依存リンク多重化方式の構成

平成 26 年度には、本方式を実装した試作機（ポートエキスパンダ装置）を開発し、多重伝送実験を行い提案方式の有効性を実証した。ポートエキスパンダ装置には、10G/40G/100Gbps IEEE802.3 規格用 (SNAP12、CFP、CFP2)、8G/16G FibreChannel 用 (SFP+、SNAP12)、Infiniband 用 (SNAP12) を収容するインタフェースを搭載した (図 2)。装置内部には、これらインタフェースを多重化する FPGA、及び 400Gbps に多重化した光インタフェース (CFP2 を 4 台使用) を搭載した (図 3)。

また、他の課題 3-I (b) 「光多値向け高性能信号処理技術」、3-I (d) 「マルチコアファイバ光接続技術」とも連携した合同の実証実験を行った。特に多重化後の 400Gbps 高速リンクとして課題 3-I (b) の 400Gbps 光インタフェースを用い、課題 3-I (d) によるマルチコアファイバを用いたネットワークを用いた連携動作を確認した。

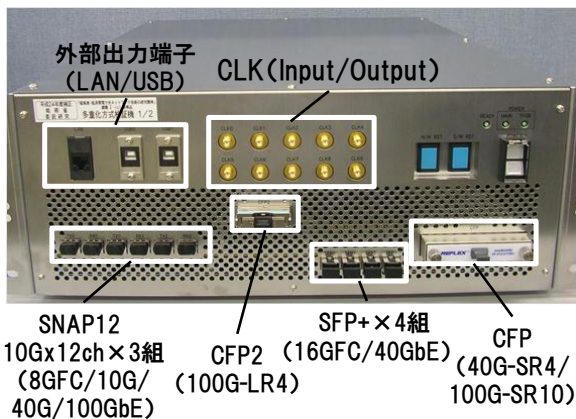


図 2: ポートエキスパンダ装置 (フロントパネル)

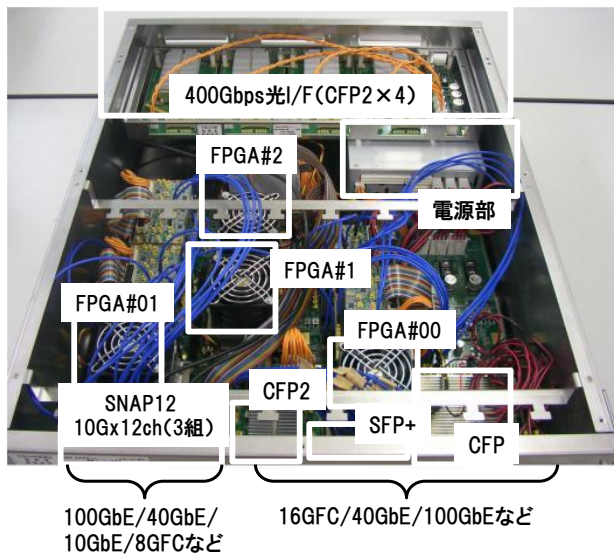


図 3: ポートエキスパンダ装置 (内部)

課題ア) 物理リンク多重化技術

本サブ課題の到達目標は、2種以上の異なるリンク層プロトコル、2種以上の異なる伝送レートの低速リンク (8本以上) 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認することである。また、多重化処理においては 1 マイクロ秒以下 (数百ナノ秒オーダを目標) の処理遅延で実現することを目指した。

本技術では、物理リンクで伝送するビットデータを、デ

ータフォーマットや伝送速度には手を加えず、トランスペアレントに多重・分離する伝送手法を用いた。リンク多重化側では、ユーザからの低速の物理リンク (低速リンク) を収容し、そこで受信したビットデータを一旦固定長のブロックに分割する。この固定長ブロックの一連のフローを仮想レーンとして扱う。仮想レーン 1 本は固定速度 (試作では 6.25Gbps) とし、それを低速リンクの伝送レートに包含する必要本数分束ねたマルチレーンの構造にすることで、様々な伝送レートに対応する。更に、他の低速リンクから生成された仮想レーンを束ね、高速の多重化リンクを構成する (図 4)

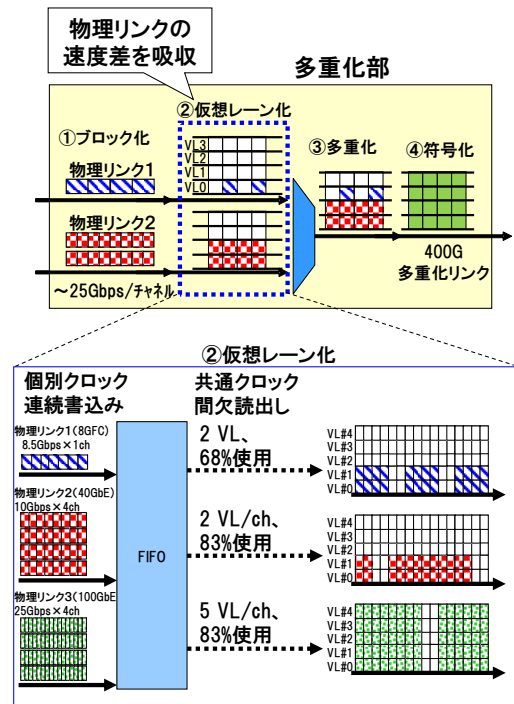


図 4: 多重化部の構造

本サブ課題の最終目標であるプロトコルおよび伝送レートに無依存であり、多数の低速リンクを論理的に束ねて 400Gbps の大容量リンクに重畳するという多重化技術の確立に向けて、平成 24 年度には要素技術の個別検討から開始した。具体的には、課題となる低速リンクから仮想レーンへの変換方法、その際のオーバーヘッド低減方法や、高効率な仮想レーン多重方法、異速度リンクの多重にも対応した仮想レーンへの割当方法について、個別の検討を実施した。また一部の機能については、論理回路試作を実施した。検討した方式および論理回路は、論理シミュレーション上での動作確認だけでなく、方式検証を行う試作機 (多重化方式検証機) を開発し、実速度の試作機上で動作検証も行った。多重化方式検証機は多重・分離機能の論理回路検証を行う目的で開発したものであり、多重化部 FPGA および低速リンクの光モジュールのみで構成されている。これは丁度、最終年度に開発したポートエキスパンダ装置の多重化部 FPGA と低速リンク部を切り出した構成に等しい。

平成 25 年度には、個別に開発した機能ブロックを結合し、全機能ブロックの統合論理回路シミュレーションを実施し、期待する動作を得られた。その際、論理回路シミュレーションにより多重化伝送 (多重化処理および分離化処理) の遅延時間を検証し、 $0.82 \mu\text{s}$ となることを確認した。これにより多重化処理 $1 \mu\text{s}$ 以下、という本開発の目標を達成した (図 5)。

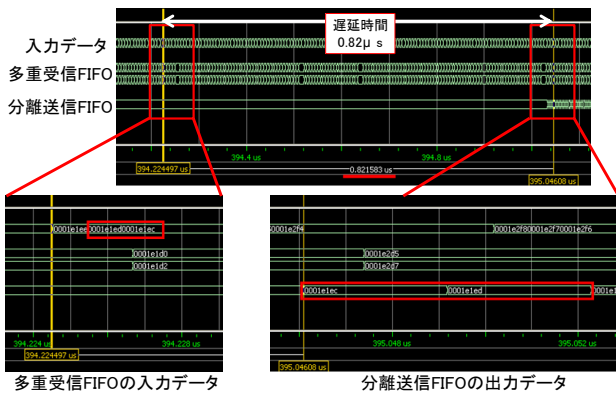


図 5：多重化処理の論理回路シミュレーション結果

課題イ) 伝送レート透過・再生技術

本サブ課題の到達目標は、2種以上の異なるリンク層プロトコル、2種以上の異なる伝送レートのリンクの多重化において、異なる源信クロックから生成された非同期リンクに対応することで、プロトコル無依存、かつ伝送レート無依存を実現するために必要な、伝送レートの透過・再生技術を確認することである。

本方式の手法は、多重化装置の低速リンクで受信した伝送レートと、多重化伝送後の分離化装置で低速リンクから送信する伝送レートを保障するものである。多重化側では、低速リンクから受信したビット列データを、一旦固定長のデータブロック（固定長ブロック）に変換し、それを固定レートの仮想レーン（伝送レートによって複数本使用）に乗せ、更に多数の仮想レーンを束ねた多重化リンクを構成して、分離化装置へと送達する。

分離化側では、多重化リンクから、低速リンク毎に纏められた仮想レーンの束を抽出し、元の低速リンクのビット列データの順に固定長ブロックを受信 FIFO へ書き込む（図 6）。受信 FIFO への書き込みは、全仮想レーンで同期したクロックによって行われるが、読み出しの際は、低速リンクに対応した個別的可変クロック生成部により生成される読み出しクロックに同期して読み出す。その際、受信 FIFO でオーバー/アンダーフローが生じないように、生成クロックの周波数を増減する制御を行う。これにより±100ppmの非同期源信のリンクに対応する。

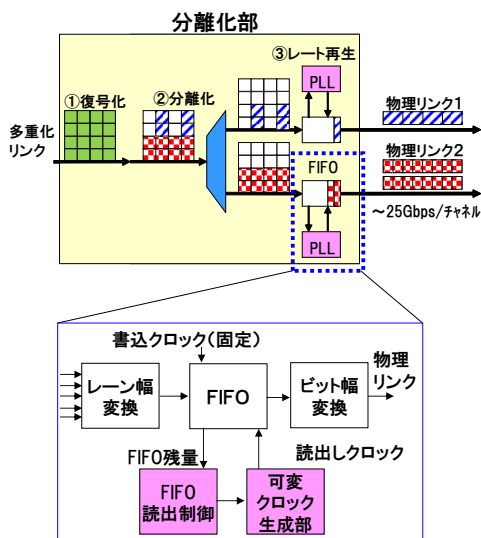


図 6：分離化部の構成

本サブ課題の平成 24 年度には、伝送レートが 5Gbps 帯から 25Gbps 帯、及び±100ppmの周波数または伝送レート誤差を持つ非同期の物理リンクからのデータを受信し、これを固定レートの伝送を通過した後も、同等な伝送レートの再生を可能とする「伝送レート透過・再生方式」の基礎検討、仕様決定を行い、一部の機能については論理回路試作も実施した。本方式では、最終段の受信 FIFO 読み出しクロックを作る可変クロック生成部として、離散的なクロックを生成する周波数シンセサイザを用いる手法を取った。

平成 25 年度は、伝送レート再生方法の安定性を向上する「DAC 制御 PLL」の検討を行った。前年度の、周波数シンセサイザを制御する方式では、再生クロックが離散的かつ変化が急峻であったため、これにより再生された低速リンクを受信する側（ユーザ側）で、同期外れが生じる不具合があったためである。改善版のクロック制御方式では、生成クロックを生じる VCO の制御電圧を、DAC 出力のアナログ電圧で直接制御することとした。DAC のアナログ電圧出力も離散値だが、時定数により緩やかに変化するようにした（図 7）。これにより、VCO 出力クロックの中心周波数は離散的ではあるが、離散値間を緩やかに変化するため、受信側同期（CDR）が追従できるようになった。

DAC 制御 PLL を含め、試作したポートエキスパンダ装置に実装した論理回路の評価も行った。評価実験では、多重化部で受信した伝送レートから参照クロックを検出し、分離化部では DAC 制御 PLL の出力周波数を設定出来ることを確認できた。また、周波数誤差±100ppm以下の非同期クロックにも対応する目標については、±400ppmまで対応するなど、上回って達成した（図 7）。これは、IEEE802.3規格であれば±100ppmの仕様で充分だが、他の規格ではそれ以上の性能が必要であるため、目標仕様を拡大した結果である。

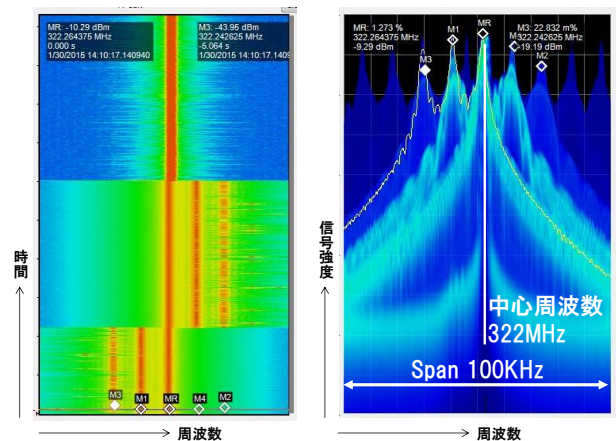


図 7：非同期源信への追従（±400ppm）

課題ウ) 物理リンク管理技術

本サブ課題の到達目標は、2種以上の異なるリンク層プロトコル、2種以上の異なる伝送レートのリンクの多重化において、多重装置と分離装置のポート割当の一貫性を維持可能な、物理リンク管理技術を確認することである。

本サブ課題では、上位層プロトコルで扱う宛先情報等に依らないリンク管理方式の開発を実施した。本方式では、仮想レーン単位に、多重装置と分離装置のポート割当てを管理し、管理情報を定期的に多重装置一分離装置間でネゴシエーションすることで、ポートマッピング対応の一貫性を維持する（図 8）。

断発生時の影響を確認した (図 10)。

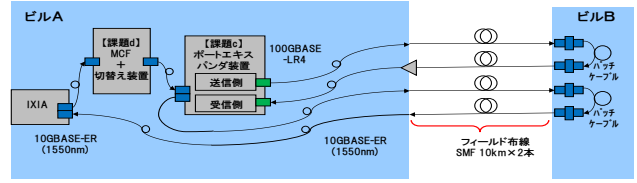


図 10: 実証実験の接続構成

3. 今後の研究開発成果の展開及び波及効果創出への取り組み

研究開発の進捗に合わせて、国内外において成果発表及び標準化活動等を実施し、本研究成果のアピールを行った (図 6)。標準化活動については、特に物理リンク多重化技術は OIF PLL MLG (Multi-link Gearbox IA 等) の調査、400Gbps の多重化及び符号化技術は IEEE 802.3 400GbE SG (Study Group) の調査、及び提案 (平成 25 年 7 月) を行った (図 11)。また IEEE P802.3bs 400GbE TF (Task Force) における標準化状況や検討フェーズ、技術動向に関して調査を行った。特に IEEE 802.3 への提案活動については、実施期間中の検討フェーズが光伝送方式に関する部分にフォーカスされ、多重化技術としての提案が困難であった。そこで本プロジェクトの課題 I-b 「光多値伝送向け高性能信号処理技術」のメンバと協力して提案を行った。

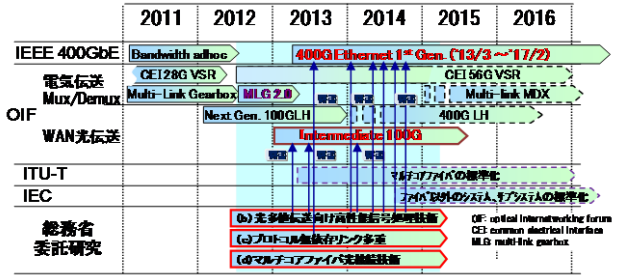


図 11: 標準化への対応状況

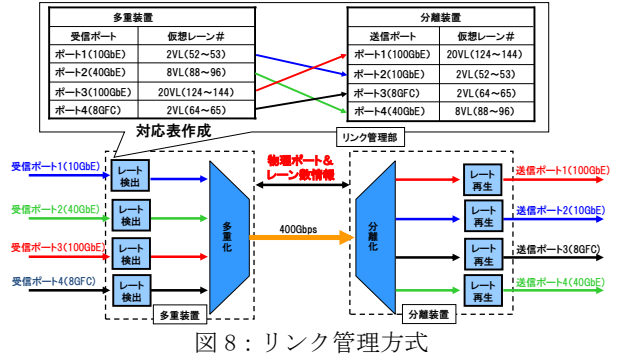


図 8: リンク管理方式

本サブ課題の平成 24 年度は、低速度の物理リンクと複数本の固定レートの仮想レーンとのマッピング方式、及び受信ポート (多重装置) と送信ポート (分離装置) 間でのポートマッピング同期方式の基礎検討を実施した。平成 25 年度はこれを基に、主要機能の詳細機能ブロックへのブレイクダウン、各機能ブロックの動作詳細化を実施した。また、平成 24 年度にて要素開発したリンク多重化方式論理回路部に対し、物理リンク管理方式機能部位の論理回路を更に結合する論理回路部位の開発を実施した。また開発・製造したポートエキスパンダ装置を用いて、3つのサブ課題による提案方式の連携動作する評価実験を行った。これら論理回路シミュレーション、及び実機評価及びにより、低速リンクから仮想リンクへのマッピングを含むリンク管理動作を確認した。

平成 26 年度には、過年度に開発したポートエキスパンダ装置の FPGA 上に、物理リンク多重化方式と伝送レート透過・再生方式、及び物理リンク管理方式を統合した論理回路を実装した。また、物理リンク管理方式では、複数台のポートエキスパンダ装置間を接続する際のリンク管理に関する開発を行い、論理回路及びソフトウェアをポートエキスパンダ装置内部の FPGA へ追加で実装した。同装置を用いた評価実験により、最大 8 ポートまでのリンク多重化の確認までを確認した (図 9)。

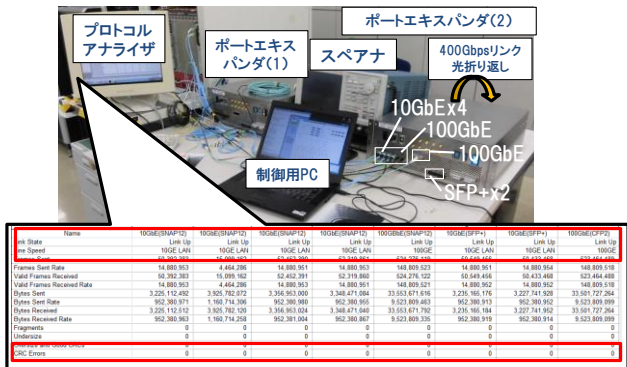


図 9: 8 ポート同時接続での評価結果

フィールド実証実験

平成 26 年度には、プロトコル無依存リンク多重化方式を実装したポートエキスパンダ装置を用いて、フィールドに敷設された光ファイバでの統合実証実験 (本プロジェクトの他課題との連携接続実験を含む) を実施した。実証実験は、札幌にある複数のビルを光ファイバで結び、総伝送距離 20km の加入者系光ネットワークを模擬したテストベッドを構築して行った。本実証実験により、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、及び伝送経路が災害等により断線した際のリンク瞬

IEEE P802.3bs は、現状も議論が継続中であるため、本プロジェクト終了後も参加を継続し、必要に応じて提案活動を実施していく。OIF については、MLG (Multi-lane Gearbox) Ver. 2.0 が策定された後、平成 27 年以降より「FlexEthernet」という新たな多重化方式の議論が開始している。FlexEthernet では、複数の低速リンクを、1 本の高速 PHY へ多重することなどが議論されており、本研究の成果適用の一つとして今後検討していく。

実用化に向けた展開としては、IEEE P802.3bs の標準化完了が平成 30 年と予定されており、実用化及び最初の市場展開は平成 30 年以降と推定される (図 12)。現行最速の 100GbE がサーバ市場でも普及し、それをアグリゲーションするネットワークでは次世代高速インタフェースが期待される時期であるため、本研究の実用化による波及効果は大きい。

これに向けて通信装置およびソリューション製品へ、400Gbps インタフェース及びプロトコル無依存リンク多重化方式の適用に向けた技術開発を検討していく。

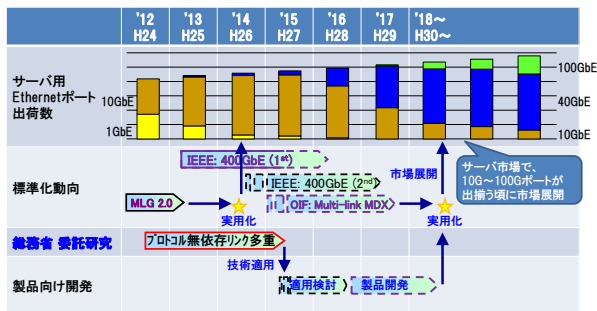


図 12： LAN/MAN 市場動向と実用化時期の予測

4. むすび

本研究課題では、次世代の 400Gbps インタフェースが普及する時期に向け、異なるフォーマットのプロトコルを載せた多数の低速インタフェースを、高速インタフェースに多重化して伝送するプロトコル無依存リンク多重化技術の研究を行い、また当技術を適用したポートエキスパンダ装置を試作して動作実証を行った。

【査読付発表論文リスト】

- [1] 光野正志、堤聡、豊田英弘、“プロトコル無依存リンク多重化技術の開発”、電子情報通信学会論文誌 B Vol.J97-B, No.10, pp.973-976 (平成 26 年 10 月 1 日)
- [2] 光野正志、堤聡、豊田英弘、“Development of Port Expander Using Protocol-independent Link Aggregation Technologies for Large-Scale Data Center Networks”、APCC2014、T1D-3 (平成 26 年 10 月 2 日)

【国際標準提案・獲得リスト】

- [1] IEEE802.3・July 13 Plenary Meeting、“Initial Thought about Modulation Format & FEC for "Long-reach" 400GbE”、提案 平成 25 年 7 月 16 日

【報道掲載リスト】

- [1] “データ通信速度 4 倍に 日立製作所”、日本経済新聞、平成 27 年 3 月 10 日
- [2] “毎秒 400 ギガビット高速通信技術 日立、データセンタ向け”、日経産業新聞、発表 平成 27 年 3 月 10 日
- [3] “大規模分散型データセンタ向け 日立が大容量・高信頼伝送技術”、電波新聞、平成 27 年 3 月 10 日

【本研究開発課題を掲載したホームページ】

<http://www.hitachi.co.jp/New/cnews/month/2015/03/0309.html>