

IOWNを支えるディスアグリゲータッドコンピューティング

IOWNの実現のためには、これまでと比較にならない大量のデータを効率良く処理できる高度なコンピュータが求められます。NTTでは、この要求にこたえるために、光電融合技術を最大限活用した画期的なコンピュータアーキテクチャについて検討を行っています。本稿では、この新たなコンピュータアーキテクチャである「ディスアグリゲータッドコンピューティング」の全体概要と基本コンセプトについて説明します。

おかだ
岡田

あきら
顕^{†1}

きはら
木原

せいじ
誠司^{†2}

おかざき
岡崎

よしかつ
義勝^{†3}

NTT先端集積デバイス研究所 所長^{†1}

NTTソフトウェアイノベーションセンタ 所長^{†2}

NTTネットワークサービスシステム研究所 所長^{†3}

IOWNが求める コンピューティング能力

あらゆるものがネットワークにつながり、社会・経済・文化のあらゆる活動にICTによる変革がもたらされつつある近年においては、数々のデータを相互につなげることにより、個々のデータだけでは得られない新たな価値を生み出していくことが重要となっています。このようなスマートな社会の実現のために、高速にデータを転送する広帯域ネットワークや、これまでにない高い処理能力を持つ情報処理装置が求められます。

NTTではこのようなスマートな社会を実現するために、IOWN (Innovative Optical and Wireless Network) 構想を提案しています⁽¹⁾。また、IOWNの実現に向けて、想定されるユースケースを議論し、これらユースケースを実現する主要技術の検討を進めるための仲間を募り、IOWN Global Forumを設立しました⁽²⁾。IOWNは、それがもたらす世界観とサービス、さらには

それを実現するためのネットワークおよびシステムまでを含む広大な構想です。

これらIOWNの世界を実現する技術の構成を図1に示します。高速・低遅延に光を最大限活用したデータ伝送を行うオールフォトニクス・ネットワーク (APN)、大量なセンシングデータを基にデジタル空間に実世界を再現するデジタルツインコンピューティング (DTC)、伝送からアプリのレイヤまで統合的に制御を行うコグニティブ・ファウンデーション (CF) という3つの技術分野を定義し、それぞれの検討を始めています。

これらIOWNを支えるAPN、DTC、CFのすべての技術分野に求められるのが、高い処理能力を持つコンピュータです。DTCでは、膨大なセンサ等のデータ源から、これまでと比較にならない精度・粒度の大量のデータを吸い上げ、それらを処理・加工するために、現状から桁のオーダーで高いコンピュータの処理能力が求められます。またルータや携帯基地局のような、専用装

置で実現されていた機能も含めてソフトウェアで実現されるIOWNの時代になると、高いバケット処理能力やスケジューリング能力が求められるAPNを構成する各ノードにおいても、高いコンピューティング能力が必要となります。さらにCFでは、コンピューティングリソースを適切に制御し、さらにコンピューティングリソースとネットワークを統合的に制御するという、これまでになかった新たな制御・マネジメント能力を提供する必要があります。そのためには大量のネットワークやコンピュータへの要求を集約・分析し、適切な帯域・波長、さらにはコンピューティングリソースを高度なリアルタイム性で割り当てる、極めて高い処理能力が求められます。従来のコンピュータアーキテクチャでこの高い処理能力を実現するためには、大量のコンピュータが必要となり、電力消費が著しく増加してしまいます。IOWNの実現にあたっては、環境負荷を高めることなく所望の高い処理能力が得られるよう、電力当りの処理能力が極めて

Innovative Optical and Wireless Network (IOWN:アイオン)

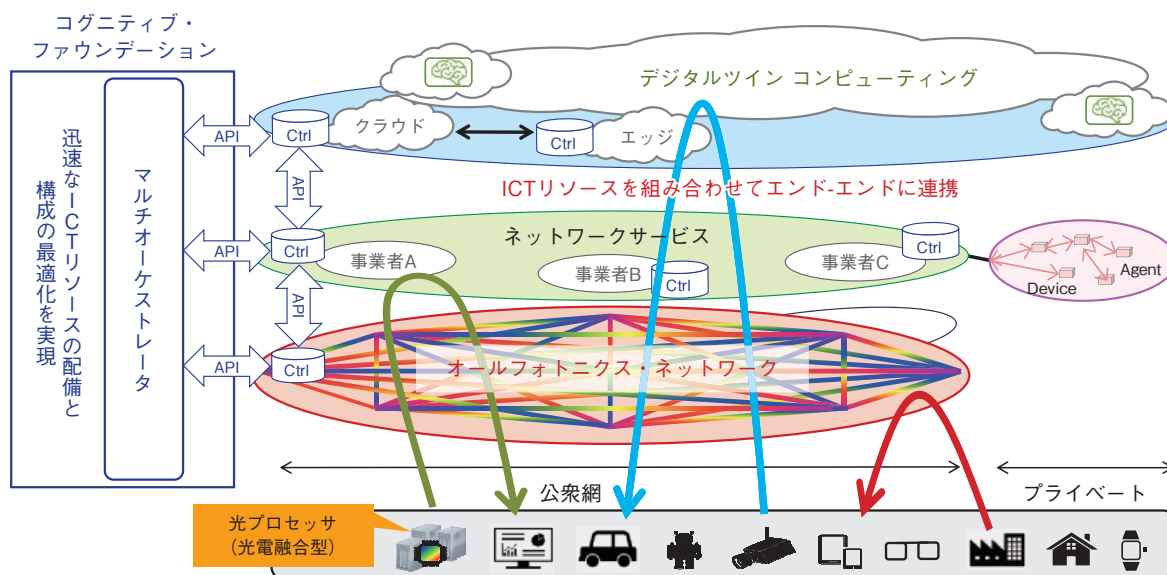


図1 IOWNの全体像

高いコンピュータが必須となります。NTTではこの課題を解決するため、光電融合技術を活用した新しい「ディスアグリゲータッドコンピューティング」という概念を提案しました。

現在、NTT先端集積デバイス研究所（先デ研）、NTTソフトウェアイノベーションセンタ（SIC）、NTTネットワークサービスシステム研究所（NS研）の3つの研究所が連携し、このディスアグリゲータッドコンピューティングの実現に向けたコンピュータアーキテクチャの研究を始めています。

IOWNを支えるディスアグリゲータッドコンピューティング

「ディスアグリゲータッドコンピューティング」と名付けた新しいアーキテクチャの概念を図2に示します。新しいアーキテクチャは、光の持つ高速性・低消費電力性・低損失性を最大限引き出す物理構成（ハードウェアアーキテ

クチャ）、論理構成（ソフトウェアアーキテクチャ）、制御方式を組み合わせ、これまでのコンピュータと異なる圧倒的な高い性能をねらうものです。箱に閉じた「コンピュータ」をネットワークでつなぐこれまでのサーバオリエンテッドな概念から、CPU（Central Processing Unit）やメモリ等のリソースを直接光で接続、光の高速性・優れた伝送特性を活かし、ラックスケールやデータセンタスケールの1つのコンピュータとして扱う概念へのパラダイムシフトです。

その基礎を支えるのは、エレクトロニクスからフォトリクスへの流れの中、電気インタフェースでは不可能だった大容量・長距離・低消費電力化を実現する光電融合技術です。

図2(a)に示す従来の物理サーバオリエンテッドな概念では、サーバ内で完結しない処理を実行するために相互のサーバ間でデータのやり取りが必要

になった際には、外部ネットワークでデータ送受を行う必要があります。コンピュータ間を接続するネットワークは、汎用的なデータ伝送手段としてリーチャビリティ確保、パスやセッション管理等、ネットワークやサービス機能提供のための深い階層のプロトコルスタックを持つよう進化してきたため、サーバ間の通信はプロトコルスタック処理のオーバーヘッドが大きくなっています。

これに対し、図2(b)のディスアグリゲータッドコンピューティングでは、光電融合技術を利用したインターコネクト（Photonic Fabric）を介して数多くのCPUやGPU（Graphics Processing Unit）、FPGA（Field Programmable Gate Array）などの演算リソースが接続されています。このためリソース間の通信にプロトコル変換を行う必要がなく、従来アーキテクチャでは避けられなかった通信オー

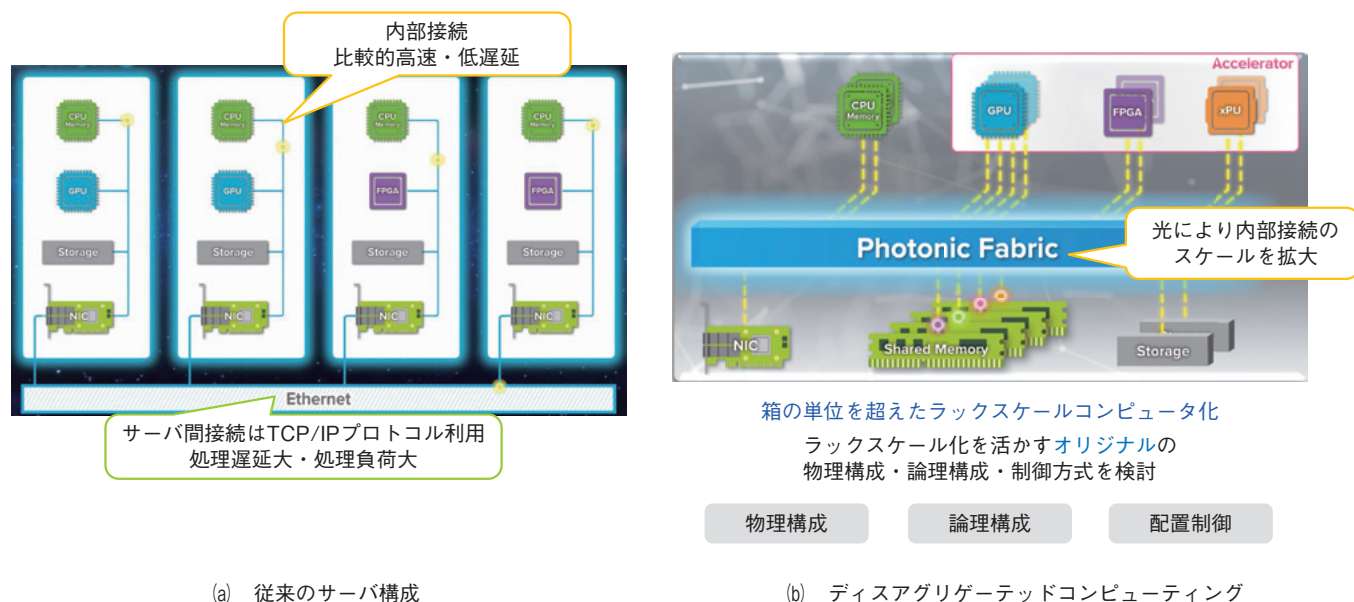


図2 ディ disaggregated コンピューティングの概念

バヘッドを大きく削減できます。さらに従来構成ではサーバ単位で電力制御やハードウェアの追加を行っていたのに対し、ディ disaggregated コンピューティングにおいては、より細かいリソース単位で電力制御・ハードウェア追加が行えるようになるため、より柔軟性の高いコンピューティング環境の提供が可能となると考えられます。

今回提案するディ disaggregated コンピューティングの3つのキーポイントのうち、物理構成を先デ研、論理構成をSIC、制御をNS研が中心となり、相互に密に協力しつつ新たなアーキテクチャの追求を進めています。

次に、各キーポイントの詳細を説明します。

■物理構成

ディ disaggregated コンピューティングの物理構成の要点を図3に示します。

電気信号では高周波になるほど伝送

路での減衰が大きくなるという物理原則のため、1本の信号線で100 Gbit/sを超えるような高速信号においては、大きな電力を消費する高出力ドライバ回路や、減衰で劣化した信号を補償する複雑な回路が必要となり、それらを用いたとしても数10 cm程度しか伝送できないという問題があります。これに対し光信号では、光ファイバで1 km伝送しても光のパワーの損失はわずか0.2 dB (4.5%) 程度であり、高速な信号を電気よりも遠くまで飛ばせるという利点を持ちます。

そのため、ディ disaggregated コンピューティングにおいては、LSI (大規模集積回路) の真横に高密度、広帯域、低消費電力の小型光電融合デバイスを搭載、LSIからの電気出力をすぐ光信号に変換する構成とします。これによりLSIからの電気信号は隣接する光電融合デバイスまでの最短距離に抑えることができ、そこから先の伝送は低消費電力の光電融合デバイスが

担うことで、LSIの高速I/O (Input/Output) の消費電力を大きく削減しつつ、コンピュータ内部インタフェースの広帯域化を実現することができます。先デ研ではこの用途に向けた光電融合デバイスの研究開発を進めており、その詳細を本特集記事『ディ disaggregated コンピューティングのための光電融合技術』⁽³⁾にて説明します。

光電融合デバイスを用いたディ disaggregated コンピュータの全体像をイメージしたのが、図4のモックアップです。アクセラレータやCPUなどを載せたカードが光配線を備えたバックプレーンで相互に接続されており、各カード上では小型・高密度な光電融合デバイスによりLSI直近で光化され、光コネクタを介しバックプレーンに接続されています。

このような構成によって、これまでの電気信号が届く箱の単位を超えたコンピュータのラックスケールの大規模

化が可能となり、カード単位での機能追加などフレキシブルかつ効率的な運用が可能になります。

■論理構成

ディスアグリゲートドコンピューティングの論理構成を図5に示しま

す。CPUは汎用的であらゆる処理をフレキシブルに実行できる利点はあるものの、その汎用性ゆえに特定処理に特化したアクセラレータと比較すると電力効率が劣るという欠点があります。そのため、ディスアグリゲートドコンピューティングにおいてはCPUへの依存率低減をねらい、特定用途に特化したアクセラレータを準備し、頻度が高い処理はこのアクセラレータで実行するかたちで電力効率を高める構成を考えています。光接続により従来のサーバより大規模化できるディスアグリゲートドコンピュータにおいては、コンピュータ全体で高頻度に行われる処理を効率良くアクセラレータに集約できるため、数多くの異なる処理に特化した複数の異なるアクセラレータを備える「ヘテロなアクセラレータプール」が実現でき、これにより電力効率を高められます。

ただ、この場合でもアクセラレータ間でのデータのやり取りにCPUの介入が必要だと、結果的にCPU負荷が高まりアクセラレータに処理をオフロードする効果がなくなってしまいま

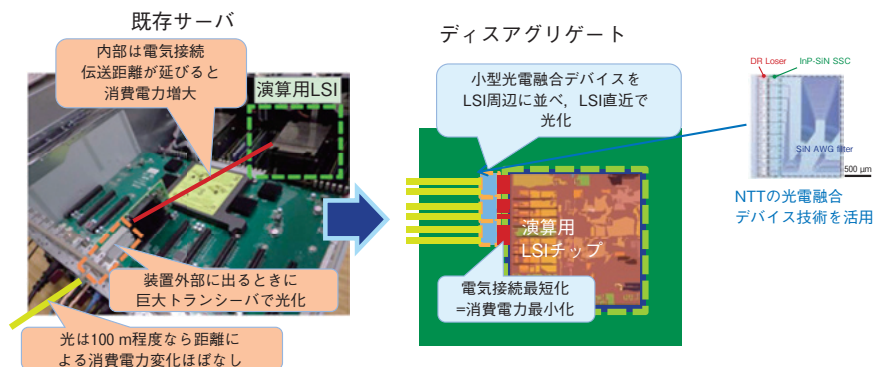


図3 ディスアグリゲートドコンピューティングのキー技術1…物理構成の概要



図4 ディスアグリゲートドコンピュータのモックアップ

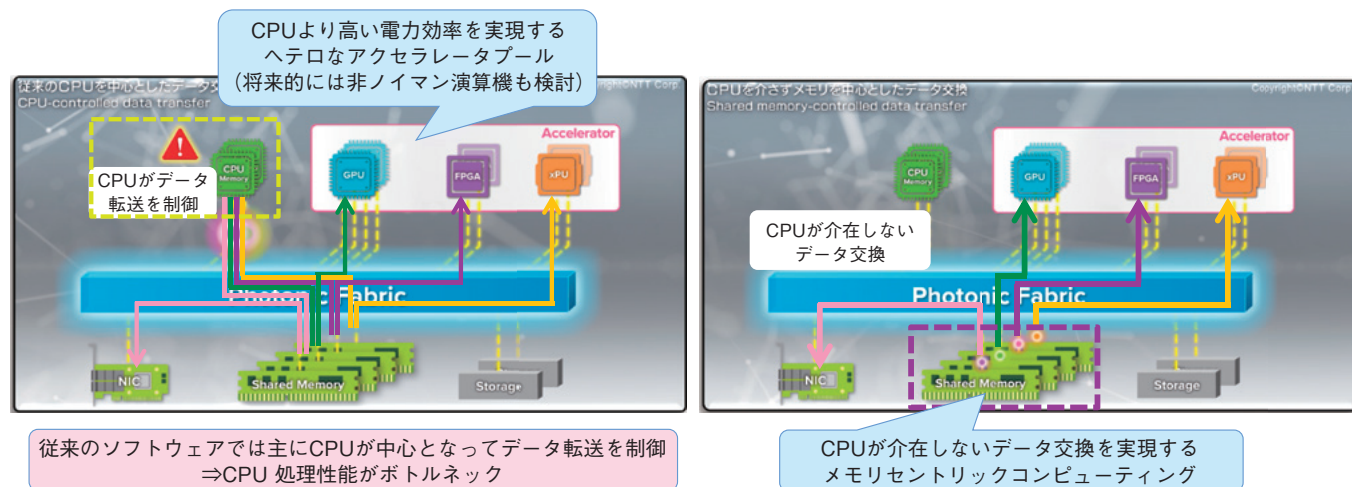


図5 ディスアグリゲートドコンピューティングのキー技術2…論理構成の概要

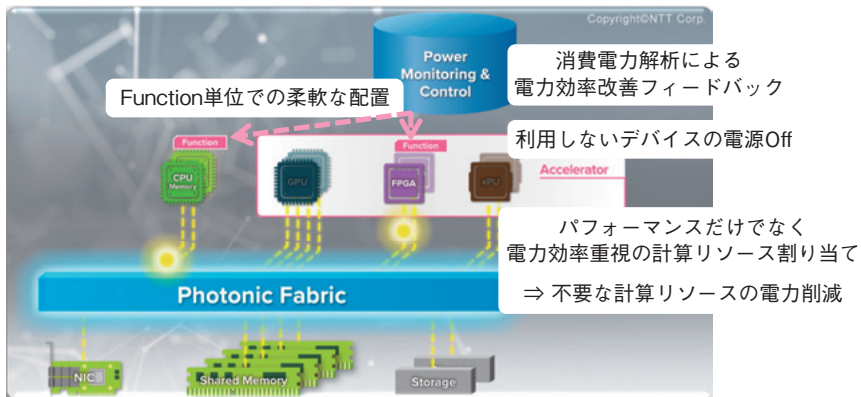


図6 ディスアグリゲートドコンピューティングのキー技術3…配置制御の概要

す。そのため、CPUの介在なしにデータ交換を行うメモリセントリック化によって、CPU負荷を大きく下げる検討をSICが中心となり進めています。このアクセラレータプール化と、アクセラレータ間のデータ交換のメモリセントリック化により大きくCPU処理を削減し、電力効率の改善をねらいます。詳細は本特集記事『ディスアグリゲートドコンピュータに向けたメモリセントリックアーキテクチャ』⁽⁴⁾にて説明します。

■配置制御

これら物理・論理構成を備えたディスアグリゲートドコンピュータを活かすためには、このアーキテクチャを最大限活かすようにソフトウェアを配置制御することが重要になります。そのため、NS研が中心となり、ソフトウェアの特性に基づいてCPUや各種アクセラレータ等の計算リソースの最適な活用をねらう「パワーウェア動的配置制御」を提案しています。

この概念図を図6に示しますが、ソフトウェアを細かなファンクション単位に分割し、消費電力を最小化するように各ファンクションで利用するアクセラレータ・CPU等のデバイスを動

的に選択し、最適なリソースにソフトウェア機能を柔軟に割り当てることで電力効率を重視した計算リソース割当てを実施します。さらには、ノーマリオフデバイスを利用して必要なときだけ電源を入れるイベントドリブ的な制御も行います。詳細は本特集記事『光ディスアグリゲートドコンピュータにおいて電力効率最大化を実現するパワーウェア動的配置制御技術』⁽⁵⁾にて説明します。

今後の展開

NTT研究所では、これまで述べたような画期的なアーキテクチャに基づくディスアグリゲートドコンピューティングを提案しています。今後は、画像推論といった代表的なアプリケーションとの組合せによる評価機の作成と評価を進め、提案アーキテクチャの定量的な実証を行っていきます。

さらに長期的には、LSIチップ内部への光の導入や、光を演算そのものに使った光演算デバイスといった最新の研究結果も取り込み、さらに先進的なアーキテクチャへの発展も検討していきます。

■参考文献

- (1) https://www.rd.ntt/_assets/pdf/techreport/NTT_TRFSW_2020_JP_W.pdf
- (2) <https://iowngf.org/>
- (3) 坂本・佐藤・瀬川：“ディスアグリゲートドコンピューティングのための光電融合技術,” NTT技術ジャーナル, Vol. 33, No. 5, pp. 45-49, 2021.
- (4) 石崎・山部：“ディスアグリゲートドコンピュータに向けたメモリセントリックアーキテクチャ,” NTT技術ジャーナル, Vol. 33, No. 5, pp. 50-53, 2021.
- (5) 金子・藤本・岩佐：“光ディスアグリゲートドコンピュータにおいて電力効率最大化を実現するパワーウェア動的配置制御技術,” NTT技術ジャーナル, Vol. 33, No. 5, pp. 54-57, 2021.



(左から) 岡田 顕/ 木原 誠司/
岡崎 義勝

NTTの3つの総合研究所にまたがる研究所が力を合わせ、これまで光通信で培ってきたデバイス技術・制御技術・ソフトウェア技術を活かした高い電力効率を持つ画期的な情報処理基盤の実現に向けて取り組んでいます。

◆問い合わせ先

NTT先端集積デバイス研究所

TEL 046-240-2022

FAX 046-240-4328

E-mail sende-kensui@hco.ntt.co.jp