

超高速TCP/IP通信ハードウェア処理エンジン NPEngine™

NPEngine™ Ultrahigh-Speed Hardware-Based TCP/IP Processing Engine

田中 信吾 山浦 隆博

■ TANAKA Shingo ■ YAMAURA Takahiro

近年、映像の高画質・高解像度化により、ネットワーク上を流れる映像や音声などのリッチコンテンツが増加している。ネットワークの通信プロトコルには、主にTCP/IP (Transmission Control Protocol/Internet Protocol) が使われているが、現在その処理は機器に組み込まれたCPUによって行われており、データ伝送帯域の増大に伴うCPUの動作周波数や消費電力の増大が問題となってきた。

この問題を解決するため、東芝は、TCP/IPのデータパケットの処理を専用ハードウェアで行うTCP/IP通信ハードウェア処理エンジンNPEngine™を開発した。新たに開発したダイレクト転送方式及びパイプライン処理により、従来の組み込みCPUに比べ消費電力当たりで22~28倍、周波数当たりで80倍以上の伝送レートを実現した。

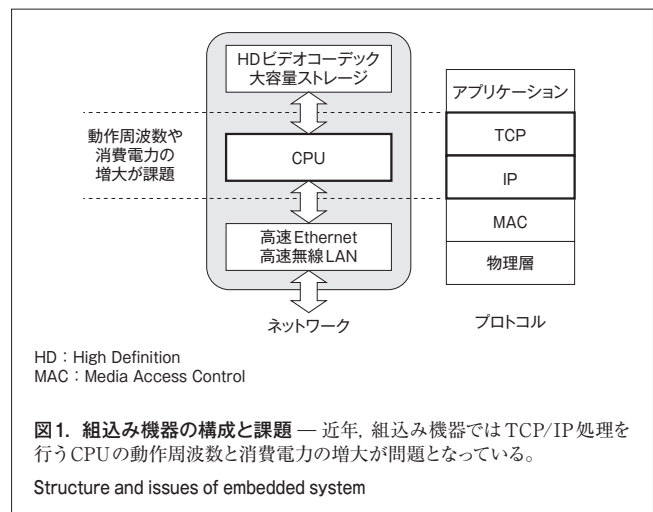
With the wide dissemination of digital high-definition (HD) broadcasting in recent years, large-volume audio and video contents have been increasingly distributed over networks including the Internet. The Transmission Control Protocol/Internet Protocol (TCP/IP), a standard network protocol of the Internet, has been conventionally processed by a software-based engine running on the central processing unit (CPU) embedded in network devices. The increase in transmission bandwidth is therefore an important issue, with CPUs requiring a higher operating frequency and higher power consumption to handle large volumes of data traffic including video streams.

To overcome this problem, Toshiba has developed the NPEngine™ hardware-based network protocol processing engine dedicated to TCP/IP, incorporating several technologies including hybrid protocol processing, direct data transfer, and pipeline processing. Experiments on NPEngine™ confirmed that it achieves much better performance than existing software-based solutions.

1 まえがき

近年、地上デジタル放送の普及に伴う映像の高画質・高解像度化が進んでおり、PC (パソコン) や携帯端末などの性能向上もありネットワーク上を流れる映像や音声などのデータ量が増え、ネットワークに要求される伝送帯域が飛躍的に増大している。ネットワークの下位層である物理層やデータリンク層では、Gigabit Ethernet, 10 Gigabit Ethernet, IEEE 802.11n (電気電子技術者協会規格802.11n) に準拠した高速無線LANなどの出現でもわかるように、伝送帯域の向上が着実に進んでいる。しかし、上位層のTCP/IPの通信プロトコル処理は依然としてCPU処理に頼っているため、特にネットワーク関連の組み込み機器の分野ではCPUの動作周波数や消費電力の増大が問題となっている (図1)。

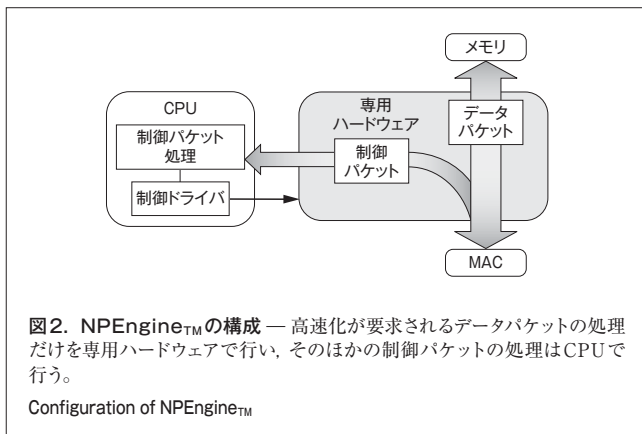
この問題を解決するため東芝は、TCP/IP通信処理のデータパケットの処理を専用ハードウェアで行うTCP/IP通信ハードウェア処理エンジンNPEngine™を開発した。ここでは、NPEngine™の概要と特長、試作品による性能評価結果について述べる。



2 NPEngine™

2.1 概要

NPEngine™の構成を図2に示す。NPEngine™は、従来すべてCPUで行っていたTCP/IP通信処理のうち、高速化が期待されるデータパケットの処理だけを専用ハードウェアで、そのほかの制御パケットの処理をCPUで行うハイブリッド構



成となっている。

TCP/IP通信では、データを転送するデータパケットのほか、デバイスアドレスを通知するARP (Address Resolution Protocol) パケット、疎通確認に用いるICMP (Internet Control Message Protocol) パケット、転送を制御するためのコネクション接続 (TCP SYNセグメント)・切断 (TCP FINセグメント) など、様々な制御パケットが送受信される。これら制御パケットに対する処理は複雑なものとなり、そのままハードウェアで実装すると回路規模が増大してしまう。一方で、これらの制御パケットはデータを転送するパケットではないため、ハードウェア化による高速化を行っても期待するデータ転送速度の向上には貢献しない。

そこでNPEngine™では、図2のようにMAC (Media Access Control) に近い段階でデータパケットと制御パケットを分離し、制御パケットは従来と同じくソフトウェアで処理を行いながら、データパケットだけを専用ハードウェアで処理する構成とした。

このような構成にすることで、ハードウェアの回路規模を小さく抑えることができ、新規開発に要するコストやデバイスコストを低減するとともに、デバイスに電源を入れただけで消費されるスタティック電力の削減を実現している。

更に、データパケットの処理を行う専用ハードウェアは、可能な限り低い動作周波数及び消費電力で高い伝送レートを実現することを目指し、ダイレクト転送とパイプライン処理という技術を採用することで実現している。以下に、それらについて述べる。

2.2 ダイレクト転送

NPEngine™の主な特長は、送受信するデータパケットをアプリケーションが確保したバッファ領域に直接読み書きするダイレクト転送方式にある。

従来のソフトウェアによる受信処理では図3(a)に示すように、まず、受信パケットはMACによっていったん一時バッファに書き込まれ、その後、TCP/IP処理を行うCPUによってアプリケーションバッファにコピーされており、これら複数回のコ

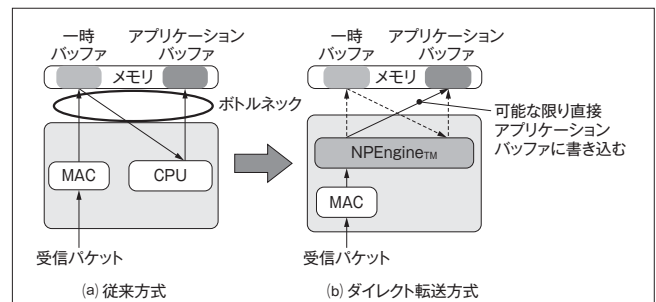


図3. 従来の受信処理とダイレクト転送方式による受信処理 — ダイレクト転送方式では直接アプリケーションバッファに対してデータの読み書きを行うことで、従来の処理でボトルネックとなっていたデータのコピー処理を削減している。

Conventional receiving processing and receiving processing of NPEngine™ with direct transfer technology

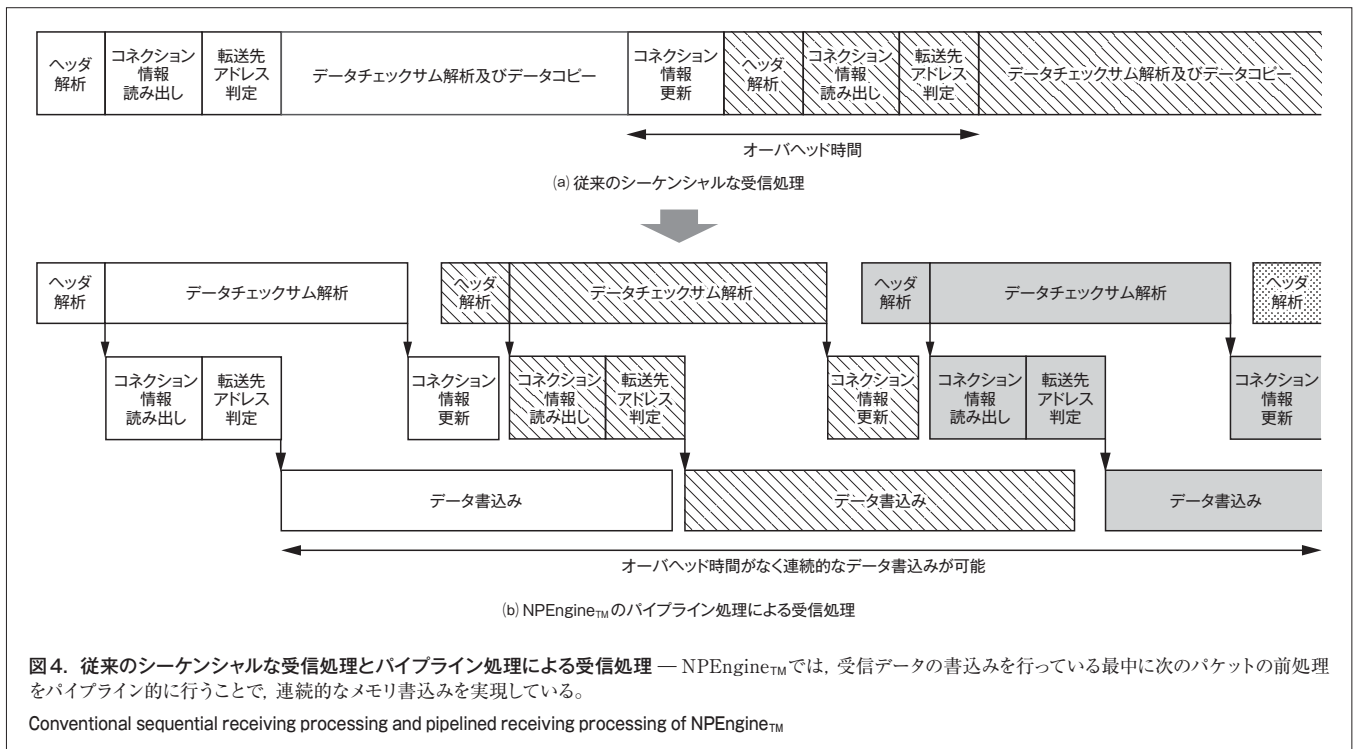
ピー処理がボトルネックとなっていた。この問題を解決するため、NPEngine™では図3(b)に示すように、MACが受け取った受信パケットをそのままNPEngine™が受け取り、メモリに書き込む前にプロトコル処理を行い、パケットに付加されているシーケンス番号からデータのメモリ上の書込み位置を特定し、可能な限りデータを直接アプリケーションバッファに書き込む。ただし、データがアプリケーションバッファからはみ出た場合や、そもそもアプリケーションバッファがまだ指定されていない場合など、データを直接書き込めない場合には、例外的に一時バッファを用いる。

このような処理をあらゆる条件下で動作可能にするダイレクト転送方式を新たに開発することで、前述のコピーによるボトルネックを解消した。また、アプリケーションが任意のタイミングで任意のバッファサイズを指定することにも対応可能で、柔軟性を損なうことなく性能向上も実現している。

2.3 パイプライン処理

NPEngine™では更に、ハードウェアの処理効率を高めるため、内部処理のパイプライン化を行っている。

従来、図4(a)に示すパイプライン化を行わないシーケンシャルな受信処理では、パケットを受信すると、まずTCP/IPのヘッダの解析を行ってパケットが属するコネクション、つまり自身のIPアドレス及びポート番号と、相手のIPアドレス及びポート番号を識別する。コネクションを識別すると、次に、現在どこまで受信したか、データを書き込むバッファのアドレスなど、そのコネクションに関連する様々な情報を読み出し、これらの情報を用いて受信データのメモリ上の書込みアドレスを判定する。アドレスを判定すると、実際にそのアドレスに対してコピーを行うと同時に、TCPデータのチェックサム計算を行う。すべてのデータの書込みが終わり、チェックサムも正しい (データが壊れていない) ことを確認すると、パケットを受信したことを記憶するため、受信したヘッダ情報に基づいてコネクション情報を更新する。このように、パケットの受信処理は複



数のステップで構成されている。

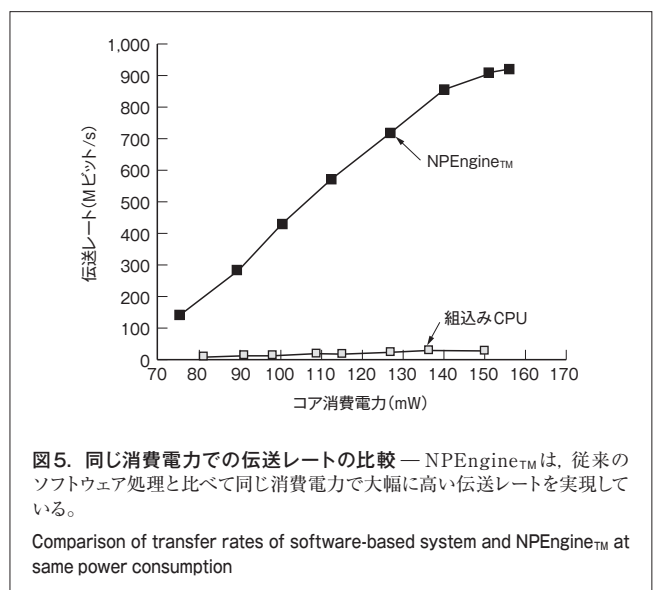
それに対してNP Engine™では、ハードウェアの並列性を生かし、これら複数のステップをパイプライン的に処理することで、このオーバーヘッド時間を解消する。まず、これらの処理系をヘッダ及びデータチェックサム解析を行うレーン（図4(b)上段）、コネクション情報の読み書きと判定処理を行うレーン（図4(b)中段）、及びデータコピーを行うレーン（図4(b)下段）の3レーンに分け、これらの処理を図のように並列動作させる。処理の順序は変えられないため、単一パケットの処理を並列化するのではなく、現在のパケットのデータ書き込みを行っている最中に次のパケットの前処理を並行して行っているところがポイントである。なお、送信処理についても受信処理と同様である。

3 性能評価

3.1 試験結果

NP Engine™の性能を確認するため、Altera社のCyclone®(注1) III FPGA (Field Programmable Gate Array) 上に試作実装してTCPのデータ受信試験を行った。FPGAにはmini-PCI (Peripheral Component Interconnect) インタフェースを設け、ネットワークインタフェースに市販のGigabit Ethernetのmini-PCIネットワークカードを用いた。従来のソフトウェアに

(注1) Altera, Cycloneは、アルテラ社の米国及びその他の国における登録商標又は商標。



よるプロトコル処理と比較するため、同じFPGA上に32ビットRISC (縮小命令セットコンピュータ) 型CPUだけを実装したものを比較対象とした。

それぞれについて、全体のクロック周波数を変化させてFPGAのコア消費電力及び伝送レートを計測した結果を図5に示す。NP Engine™はソフトウェア処理と比較して消費電力当たりで大幅に高い伝送レートを実現しており、例えば、消費電力80 mW付近では組込みCPUが約8 Mビット/sであるのに対してNP Engine™はその約22倍の200 Mビット/s、消費電力140 mW付近ではCPUが約30 Mビット/sであるのに対し

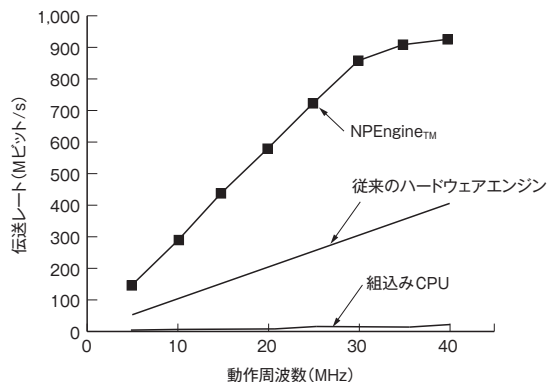


図6. 組み込みCPU及び従来のハードウェアエンジンとの伝送レートの比較
 — NP Engine™は、組み込みCPUや従来のハードウェアエンジンと比べても大幅に高い伝送レートを実現している。
 Comparison of transfer rates of embedded processor, hardware-based system, and NP Engine™

てNP Engine™はその約28倍の840 Mビット/sである。試験全体を通して、NP Engine™はソフトウェア処理の22～28倍の伝送レートを実現していることがわかる。

更に、周波数当たりの伝送レートについて、従来のハードウェアエンジンとの比較を加えたものを図6に示す。NP Engine™は、周波数当たりの伝送レートで組み込みCPUの80倍以上、特にハードウェアエンジンとの比較では3倍の性能を達成しており、この違いは主に前述のダイレクト転送方式及びパイプライン処理によるものと考えられる。

なお、ここでは受信伝送レートについての評価を述べたが、送信伝送レートについてもほぼ同様の結果が得られた。

3.2 従来技術との比較

従来技術との比較を表1に示す。

ソフトウェア処理と比較すると、回路規模は増えるものの圧倒的に高い性能を実現することができ、また、同種のハードウェアと比較すると、約1/2の回路規模で3倍近くの性能を実現することができる。更に、同種のハードウェアでは最大コネクション数に制限があるが、NP Engine™ではハードウェアとソフトウェアによるハイブリッド構成をとることで、コネクション管理を行うソフトウェアが使用可能なメモリ容量に応じて柔軟に最大コネクション数を設定することができる、というメリットもある。

表1. 従来技術との性能比較

Comparison of performances of conventional systems and NP Engine™

	従来のソフトウェア処理	従来のハードウェアエンジン	NP Engine™
追加する回路規模	なし	250 kゲート	140 kゲート
動作周波数当たりの伝送レート	0.3 (ビット/s)/Hz	10 (ビット/s)/Hz	28 (ビット/s)/Hz
最大コネクション数	無制限*	16	無制限*

*使用可能なメモリ容量に依存

4 あとがき

CPUと専用ハードウェアのハイブリッド構成によるTCP/IP通信ハードウェア処理エンジンNP Engine™を開発し、新たに開発したダイレクト転送方式及びパイプライン処理により、従来の組み込みCPUに比べ消費電力当たりで22～28倍、周波数当たりで80倍以上を達成し、また、同種のハードウェアエンジンに比べ約3倍の伝送レートを実現した。

この技術は、テレビ放送局向け番組送出サーバVIDEOS™ neoに搭載されている。この技術の適用により、サーバのコストや消費電力を低く抑えたままファイルの転送性能を大幅に向上させることが可能となり、性能の差別化に貢献している。

今後、更なる性能の向上と幅広い通信プロトコルへの対応を行い、当社のネットワーク関連機器に広く適用できる通信基盤技術として完成度を高めていく。



田中 信吾 TANAKA Shingo

研究開発センター ネットワークシステムラボラトリー研究主務。通信プロトコル処理の研究・開発に従事。電子情報通信学会会員。

Network System Lab.



山浦 隆博 YAMAURA Takahiro

研究開発センター ネットワークシステムラボラトリー。通信プロトコル処理の研究・開発に従事。電子情報通信学会会員。

Network System Lab.