URL　https://journal.ntt.co.jp/article/30140
DOI　https://doi.org/10.60249/24115005

ソフトウェア技術でネットワークの省電力化を実現する省電力イネーブラー

ネットワークを流れるトラフィックは年々増加傾向にあり，ネットワーク機器の消費する電力の増加が問題になっています．本稿では，ネットワークのソフトウェア処理において，ソフトウェアの技術により省電力化に挑戦する研究開発の取り組みについて紹介します．

キーワード：#通信ソフトウェア，#省電力，#低遅延

藤本　圭（ふじもと　けい）／名取　廣（なとり　こう）
原澤　輝（はらさわ　ひかる）／大谷　育生（おおたに　いくお）
斎藤　奨悟（さいとう　しょうご）
NTTネットワークイノベーションセンタ

ネットワーク機能のソフトウェア化と省電力

近年，専用ハードウェアで構成されてきたネットワーク装置を，汎用サーバ上にソフトウェアで構成する形態の普及が進んでいます．この実装方式は，NFV（Network Function Virtualization）と呼称されることが多く，NFVが普及した背景として，汎用CPU（Central Processing Unit）の性能向上や，通信機能のソフトウェア化技術や仮想化技術の進化の恩恵が挙げられます．NFVのメリットとして，量産された汎用サーバを使用することによりハードウェア機器の初期導入コストが安いこと，ハードウェアの回路設計等と比較してソフトウェアとして実装するため開発期間を短縮できること，動的に構成を変更させる等の柔軟な運用が可能であることなどが挙げられます．メリットが多い一方で，課題もあります．その1つが，高い性能の達成です．専用ハードウェアは，用途に応じて最適化した回路設計をするため，処理効率が高く，高い性能を達成できる傾向にあります．一方で，汎用サーバ上にソフトウェアで構成するNFVは，汎用化に伴う非効率性があるため，高い性能を達成するためには，さまざまな工夫が必要となります．その工夫として，例えばCPUを潤沢に稼動させることでパケットの到着を常時監視し即応性高く処理する，CPUが休止状態になることを抑制し休止からの復帰オーバヘッドを抑制する等が挙げられますが，これらは，いずれも省電力性を犠牲にしています．一般に，高い性能と省電力性はトレードオフの関係にあるといえます．とりわけ，高いスループットや低遅延性が求められるシステムにおいて顕著です．高いスループットや低遅延性が求められるシステムの具体的例として，vRAN（virtual Radio Access Network）が挙げられます．

ユースケースの紹介：vRAN

移動体通信ネットワークにおいて，広範なエリアへ通信サービスを提供するために，莫大な数の無線基地局が運用されています．基地局の主な役割としては，移動通信機器と無線で通信するための電波の送受信，および無線信号処理が挙げられます．この無線基地局には，高いスループットと非常に厳しい遅延要件が求められます．そのため，従来は無線基地局に最適化された専用ハードウェアが利用されてきました．近年，前述したNFVのメリットを享受するために，汎用サーバ上にソフトウェアで構成するvRANの検討が進んでいます．**図1**は

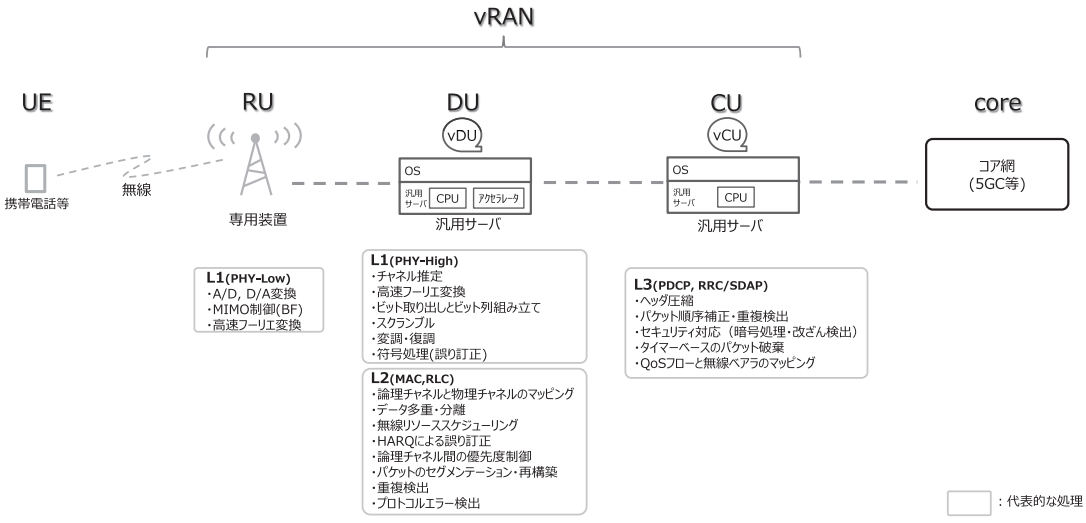

図1　vRANのシステム構成例

vRANのシステム構成例です．RANは，RU（Radio Unit），DU（Distributed Unit），CU（Central Unit）の機能部から構成されますが，汎用サーバ上にソフトウェアで実装する形態の検討が進められているのは，無線信号処理を担うDUとCUの機能部です．vRANは携帯電話等の端末との間で無線信号の送受信を行いますが，無線リソースは周波数方向と時間方向に多重化されており，マイクロ秒オーダ程度の短い間隔のタイムスロットで電波のやり取りをします．このタイムスロットに遅れることなく，DUとCUは無線信号の処理を行う必要がありますので，マイクロ秒オーダの厳しい遅延要件が求められます．また，要求されるスループットについては，例えばRUとDU間のFrontHaulと呼ばれるインタフェースにおいて，O-RAN Alliance規定のSplit 7.2の機能配備では，20 Gbit/sを超える高いスループットが要求されます．マイクロ秒オーダの遅延要件を満たしつつ，20 Gbit/sを超えるスループットを達成する必要があり，これらの機能を汎用サーバ上にソフトウェアで構成することは，技術的難易度が高いといえます．例えば，OSとしてLinuxが広く普及していますが，Linux kernelに標準搭載されたパケットの送受信機構においては，数100マイクロ秒～ミリ秒オーダの遅延が発生する場合があり[(1)]，20 Gbit/sもの高いスループットを達成することは困難です．これに加えて，厳しい遅延要件やスループットを達成しつつ，省電力化を達成することはさらに技術的難易度が高いといえます．例えば，広く普及した汎用CPUには，処理するタスクがないときは，idle状態へ遷移する機能を持っている機種が多いのですが，深いidle状態へ遷移した後に復帰するまでに要するオーバヘッドとして，100マイクロ秒を超える時間を要する場合があります．安易にidle状態へ遷移してしまうと，遅延要件に適合しない可能性があります．このように，遅延要件・スループット・省電力を同時に成立させることは，挑戦的な取り組みといえます．

ソフトウェア技術で省電力化を実現する省電力イネーブラー

NTTネットワークイノベーションセンタでは，厳しい遅延要件や高いスループットが求められるアプリケーションを対象に，ソフトウェア技術により省電力化を達成することを目標として，省電力イネーブラーPOSENA（Power Saving ENAbler）の研究開発を進めています．省電力イネーブラーの概要を**図2**に示します．ソフトウェア処理において，負荷に応じた必要最小限のコンピューティングリソースで処理するように制御することをコンセプトにしており，特に低負荷時において大きな省電力効果を実現します．

省電力イネーブラーは，複数の技術から構成されており，アプリケーション・ミドルウェア・OS・仮想化ソフトウェア等へ省電力イネーブラーをadd-in*することにより，有効化を行います．特に有望なユースケースとしては，厳しい遅延要件や高いスループットが求められるNFVアプリケーションである，vRAN，VR（Virtual Reality），AR（Augmented Reality），メディア処理，ルータ等の機能部への適用を想定しています．

本稿では，省電力イネーブラーの中から「省電力低遅延データ転送技術[(2)]」と「CPU idle考慮最適チューニング・制御技術[(3)]」の2つの技術について，有望ユースケースであるvRANでの活用例を紹介します．

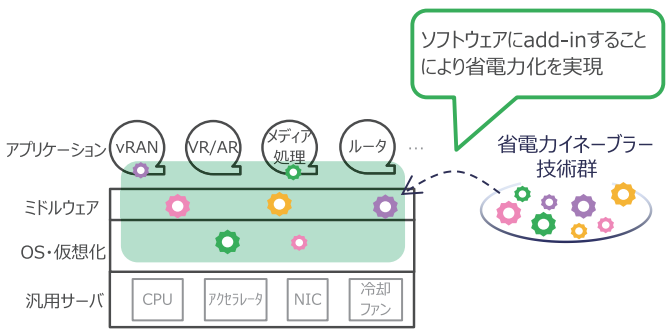

図2　省電力イネーブラーの概要

省電力低遅延データ転送技術

■従来の高速低遅延なパケット送受信方法

広く普及したOSであるLinuxには，アプリケーション開発者が容易にパケットの送受信を行うプログラムを開発することができるように，パケットの送受信を実施する機能がLinux kernel内に搭載されています．このLinux kernel内のパケット送受信機能は，汎用性を指向した設計がなされています．例えばCPUコアを1つしか搭載していないコンピュータでも動作するように，パケットが到着したときにだけCPUリソースを一時的に使用してパケット受信処理を行います．具体的には，NIC(Network Interface Card)にパケットが到着した際に，パケットが到着したことをハードウェア割込によりLinux kernelへ通知します．Linux kernelは処理していたタスクを中断し，割込を契機にパケットの受信処理を行います．このようなパケットの受信方法は，割込方式と呼称されます．Linux kernelにおける割込方式によるパケット受信は，パケットを受信したときにのみCPUリソースを使用するため，リソース使用効率が良く，CPUコアを1つしか搭載していなくても動作するため，汎用性が高い点で優れています．しかしながら，割込方式は，割込が発生すると処理中のタスクを中断し計算途中のデータを退避させる等のコンテキストスイッチによるオーバヘッドがあります．また，Linux kernel内でパケット受信処理を行うと，kernel spaceからuser spaceへパケットデータをメモリコピーする必要がある等の理由で，大きな遅延が発生する場合があり，高いスループットを出しづらいという欠点もあります．

そこで，NFVアプリケーションのように，低遅延や高いスループットが要求される場合には，Linux kernelが具備するネットワーク機能を利用せずに，kernelをバイ

＊ add-in：ソフトウェアに対し，パッチを当てる等の操作により，機能拡張を行うこと．

パスして，user spaceでパケット送受信処理を行う機構も存在します．例えば，オープンソースとして提供されているDPDK（Data Plane Development Kit）が挙げられます．DPDKは，user spaceにパケットの到着を監視するpolling threadと呼ばれるスレッドが常駐し，NICに到着するパケットを監視します．パケットの到着をCPUが常時稼動して監視するため，即応性高くパケットを受信することができます．このため，マイクロ秒オーダの遅延要件と20 Gbit/sを超える高いスループットが求められるvRANシステムにおいても，DPDKが利用されることが多いです．しかしながら，このpolling方式は，低遅延や高スループットに優れる反面，パケットの到着がない間もpolling threadがCPUリソースを常時使用してパケットの到着を監視するため，CPU使用率が常時100％になり，消費電力が高くなる問題があります．その様子を**図3** (a)に例示します．

■高速低遅延と省電力を両立するパケット送受信方法

低遅延性や高スループットを達成しつつ，省電力も同時に達成するために，前述のpolling threadに対して，図3 (b)に示すように，パケットの到着がない間はsleepさせる制御を行います．polling threadをsleepさせることにより，当該スレッドはCPUコアを常時100％使用する事態を避けることができるため，省電力化を実現することができます．しかしながら，安易にpolling threadをsleepさせてしまうと，sleepしている間に新しいパケットが到着した際，パケットの到着に気付くことができない問題が発生します．そのため，sleepしていても，即応性高くパケットの到着に気付くことができるように，あえてハードウェア割込を利用します．ハードウェア割込を発生させると，前述のとおり，割り込まれたスレッドは処理を中断しコンテキストスイッチのオーバヘッドが発生してしまうため，低遅延性や高スループット達成に向けては問題が生じます．一方で，割り込まれた際に，当該スレッドがほかに処理をしていなければ，大きな問題にはなりません．この点に着目し，パケットの到着監視とパケット受信処理を除き，ほかの処理は実施しない専用のpolling threadを設け，パケットの到着がない間は当該スレッドをsleepさせることで，sleep時にパケットが到着してハードウェア割込が発行されても，何も処理をしていない状態で割込が発生する状況をつくり出すことができます．また，ハードウェア割込は，非常に優先度が高い処理であるため，ハードウェア割込ハンドラで重い処理を実施してしまうと，システム安定動作のためのCPUタイムが捻出できず，システムが不安定になる場合があります．これを避けるために，sleep時にパケットが到着した際にNICから発行されるハードウェア割込ハンドラにおいて，sleep中のpolling threadを起床させる処理のみに制限し，パケット受信処理は，ハードウェア割込ハンドラではなく，polling threadのコンテキストで実施するようにします．このようにすることで，sleep中にパケットが到着しても，即応性高く起床しパケットの受信処理を実施できるようになります．

本技術の効果を評価するために，パケットの送受信を行う簡易的なソフトウェアスイッチへ実装し実験を行った結果，1つのpolling threadにつき，4.5 W程度の省電力効果を確認しました．sleepすることによる遅延時間の悪化は，わずか平均4～7マイクロ秒程度に抑えられていることを確認しました．また，本技術の実用性を確認するために，市中vRAN vDUプロダクトへ適用し実証実験を行い，無線信号処理に悪影響を与えずに，消費電力を削減できる効果があることを確認しました．

なお，システムに求められるスループットに依存しますが，1つのシステムに複数のpolling threadを具備する場合が多く，省電力効果はpolling threadの数が増えるごとに大きくなります．また，このsleep制御の方法は，パケット送受信処理のネットワークI/O（Input/Output）のみならず，アクセラレータI/Oやプロセス間通信等にも活用が可能です．

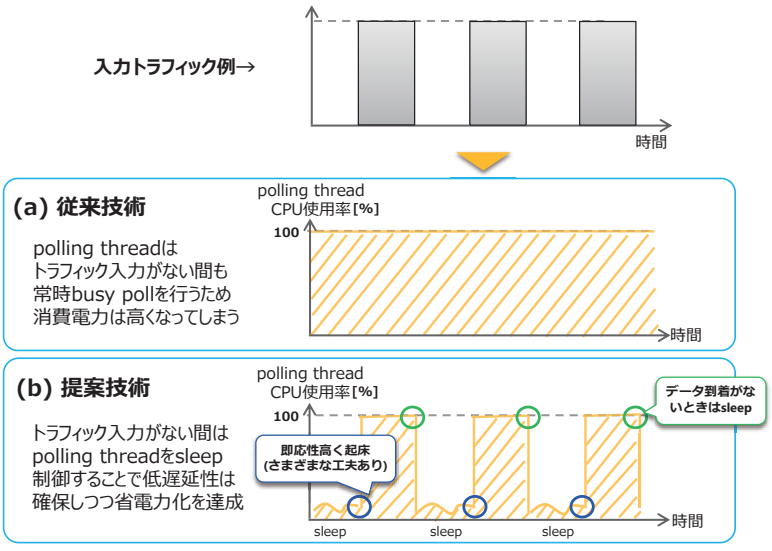

図3　polling threadによるCPU使用率の時間推移

CPU idle考慮最適チューニング・制御技術

■CPUの具備するハードウェア制御による省電力機構

広く使われているプロセッサであるCPUには，ハードウェア制御による省電力機構を具備した機種が多いです．例えば，処理すべきタスクがない間は，CPUの動作クロックを停止する，動作電圧を下げる，キャッシュ機能を無効化する等の制御を行い，省電力化を図ります．この機構はCPU idleやLPI（Low Power Idle）と呼ばれ，Intel製のCPUであればC-stateとも呼称されます．省電力化に向けては，CPU idleは大変有用な機能となりますが，一方で，CPUコアがidle状態へ遷移すると，active状態へ復帰するまでの時間が大きくなる傾向にあり，安易にCPU idle機能を使用すると，遅延要件やスループット要件を満たせない場合があります．そのため，

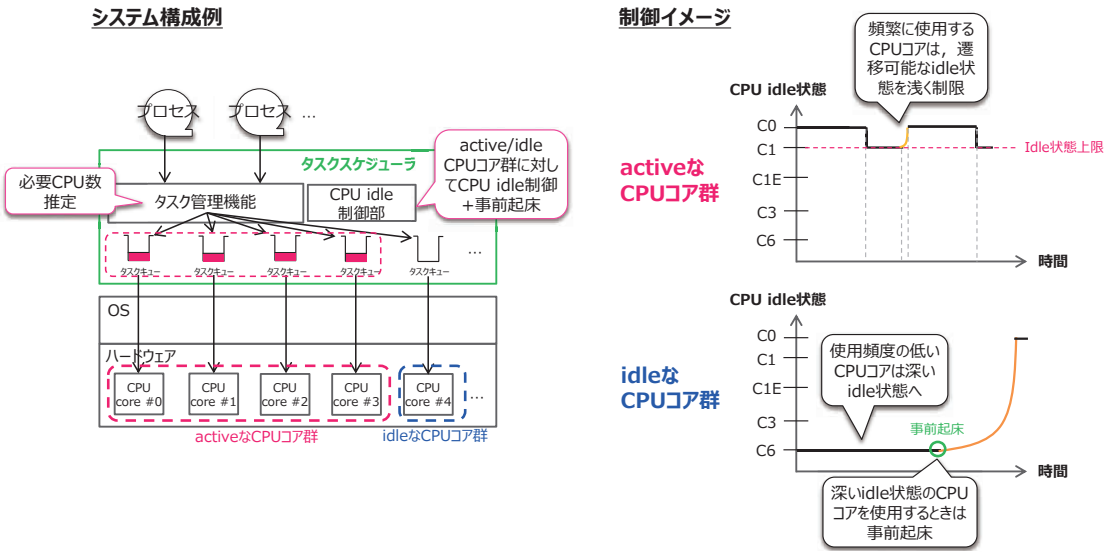

図 4　CPU idle最適チューニング・制御技術の概要

厳しい遅延要件が求められる場合や，高いスループットが求められる場合においては，CPU idle機能を無効化する運用が行われます．

■低遅延・高スループットと省電力を両立するCPU idle制御

本技術は，CPUコアに対して，深いCPU idle状態まで遷移が可能な「idleなCPUコア群」と，遷移可能なCPU idle状態を浅く制限した「activeなCPUコア群」に色分けし，これらのCPUコアへのタスクの割当てを制御することにより，省電力を達成します．概要を**図 4** に示します．activeなCPUコア群へ割り当てるタスクとしては，遅延要件が厳しいタスクや，高いスループットが求められるタスク，頻繁に処理が実行されるタスク等が望ましいといえます．対して，idleなCPUコア群へ割り当てるタスクとしては，遅延やスループットの要件が厳しくなく，処理が実行される頻度がまばらなタスク等が望ましいです．また，深いCPU idle状態へ遷移している状態のCPUコアに対してタスクを割り当てる際は，前述したとおり，active状態へ遷移するまでに大きな復帰時間を要するため，タスクを割り当てる前に，事前起床をさせる制御を行います．これにより，復帰時間の影響を受けず，深いCPU idle状態へ遷移させることによる，省電力効果の恩恵を享受することができます．

本技術の効果を評価するために，CPU idle機能を具備したCPUを搭載した汎用サーバを使用し，本技術のプロトタイプを実装し評価を行った結果，サーバハードウェアに搭載されたCPUコア数や，処理するタスクの種別等にも依存して評価結果は変動しますが，1台のサーバにつき100 Wを超える消費電力の削減効果を確認しました．また，本技術の実用性を確認するために，市中vRAN vDUプロダクトへ適用し実証実験を行い，無線信号処理に悪影響を与えずに，消費電力を削減できる効果があることを確認しました．

今後の展望

ネットワークを流れる通信トラフィックは，年々増加傾向にあります．これらの通信トラフィックを処理するためにネットワーク機器の負荷も大きくなります．そのため，技術革新がなければ，通信トラフィックの増加に伴い，消費電力も増加し続けてしまうことになります．高い性能を達成しつつ大きく省電力化を達成することは喫緊の課題であり，早期に省電力化に向けた具体的な取り組みを進める必要があります．今後も，さらに効果の高い多くの省電力化技術を創出し，実用化を進めていきます．

■参考文献

(1) K. Fujimoto, M. Kaneko, K. Matsui, and M. Akutsu: “KBP: Kernel enhancements for low-latency networking for virtual machine and container without application customization,” IEICE Transactions on Communications, Vol.E105-B, No.5, pp.522?532, May 2022.
(2) K. Natori, K. Fujimoto, and A. Shiraga: “Low-Latency and Energy-Efficient Frame Forwarding for vRAN Traffic,” NetSoft 2022, pp. 97-102, Milan, Italy, June 2022.
(3) K. Fujimoto, H. Harasawa, K. Natori, I. Otani, S. Saito, and A. Shiraga: “PWU: Pre- Wakeup for CPU Idle to Reduce Latency and Power Consumption,” SoftCOM 2022, pp. 1-6, Split, Croatia, Sept. 2022.

（上段左から）藤本　圭／名取　廣／原澤　輝
（下段左から）大谷　育生／斎藤　奨悟

増え続ける通信トラフィックに対応するためにも，ネットワーク機器の省電力化は重要な課題です．省電力化を実現する革新的な技術の創出と実用化に向けて研究開発を進めていきます．

◆問い合わせ先
NTTネットワークイノベーションセンタ
ネットワーク制御ソフトウェアプロジェクト
E-mail　nttnic-pr @ ntt.com