

令和 4 年 6 月 20 日現在

機関番号：13903

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K11940

研究課題名（和文）ネットワーク機能仮想化基盤上の通信処理の可視化と運用監視に向けた解析手法の確立

研究課題名（英文）Visualizing Network Processing in NFV Infrastructures and Developing Analytical Methods for NFV Management

研究代表者

川島 龍太（Kawashima, Ryota）

名古屋工業大学・工学（系）研究科（研究院）・准教授

研究者番号：00710328

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：5G/6Gの通信システムにおいては、要件が極端に異なる多様な通信環境を効果的に実現するため、「ソフトウェア指向」のネットワーク基盤が欠かせない。ネットワーク機能仮想化（NFV）はその中核となる概念であり、実サービスにおける導入も既に始まっている。しかし、汎用サーバを活用した具現化においては、幾重ものシステム層が内部に潜んでおり、パケット処理の挙動を精確に捉えられない。本研究では、サーバ内部におけるパケット処理過程を透明化する包括的なフレームワーク（NFV-VIPP）を開発した。具体的には、CPU使用率に頼らないシステム負荷の推定や、サーバ内部のパケット転送経路の把握などが可能になった。

研究成果の学術的意義や社会的意義

本研究課題において実現したNFV-VIPPは、（汎用サーバを基盤とした）NFVにおける実用化の課題であった、運用管理性の向上に貢献するものであり、ひいては5Gや6G通信の本格普及に一役買うような技術である。NFV-VIPPはサーバ内部におけるパケット処理の詳細を明らかにするため、そこで得られた知見を活かし、NFVシステムの基本性能向上に向けた提言も行った。性能問題はNFVの実用化に向けた最大の障壁であり、その突破に向けた道筋をつけたという点において、大きな成果であったと言える。

研究成果の概要（英文）：Software-oriented network infrastructures are bedrock for 5G/6G communications because they are based on extremely varied service requirements. Network Functions Virtualization (NFV) is the key concept for the infrastructures, and has been actually introduced in the real services. However, using commodity servers, as a means of embodiment, blinds network administrators in terms of understanding internal packet processing for management. This work addressed this management issue by making server internals transparent. Specifically, our NFV-VIPP framework enables internal system load and packet transferring path to be measurable.

研究分野：ネットワークシステム

キーワード：ネットワーク機能仮想化 システムソフトウェア 運用管理 ソフトウェア化 5G

1. 研究開始当初の背景

NFV (Network Functions Virtualisation) では、ルータやファイアウォールなどの「ネットワーク機能」は汎用サーバ上のソフトウェアシステム (NFV ノード) として実現される。NFV ノードはオペレーティングシステムや仮想化ミドルウェアなどの抽象化層を持ち、単一機器上で複数のネットワーク機能が内部接続されるという特徴がある。また、NFV ノードは DPDK (Data Plane Development Kit) などの高速パケット I/O 機構を搭載しているが、CPU やメモリなどのリソースは独自に管理される。このように、NFV ノードは特異で複雑な内部構造 (図 1) を持つため、従来のネットワーク/サーバ機器管理手法では、ノード内部で発生する通信障害の要因特定やリソース使用状況に基づくノードの状態把握ができない。

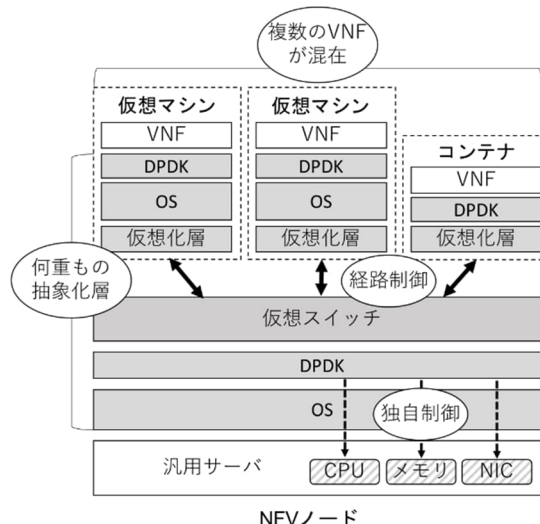


図 1 NFV ノードのシステム構成

2. 研究の目的

5G 通信のコアネットワークでは、性能要件や規模、ルーティング方針などが異なる多様な仮想ネットワークが混在するため、従来は専用ハードウェア機器として実現していた様々な「ネットワーク機能」を仮想化して柔軟なネットワーク基盤を実現する NFV 技術が必要である。従来のネットワーク機能はブラックボックスな機器として扱われ、代表的な機器管理プロトコルである SNMP や次世代版の Telemetry においても、ハードウェア部品の故障検知や CPU などのリソース使用量の監視は行われるが、機器の内部構成までは考慮しない。NetFlow や sFlow も通信フローの詳細な分析は可能であるが機器内部の状態は把握できない。しかし、NFV ノード内部のシステムアーキテクチャを考慮しない従来のネットワーク機器管理の枠組み (プロトコル、ツール、管理者の知識・経験) では、以下の要因のためにノード内部で発生するさまざまな通信障害に対処できない。

- オペレーティングシステムや仮想化ミドルウェアなどの抽象化層の存在
- 仮想マシン/コンテナ形式の複数の VNF (Virtual Network Function, VNF) による内部連携
- 仮想スイッチによる VNF 間連携のためのスイッチング処理
- DPDK による各種リソース (CPU、メモリ、NIC) の独自管理

抽象化層が増えると NFV ノード内部でのパケット処理過程やコンポーネント間通信も増えるため、パケット欠落およびキューイング遅延の発生個所や、その要因の特定が困難になる。複数の VNF が内部連携する (サービス機能チェーン) 場合は、ノード内部にもネットワークが存在するのと同義であるため、VNF 単体だけでなく、VNF 間リンクの監視も欠かせない。また、仮想マシン/コンテナ形式の VNF は動的に生成・削除されるため、内部ネットワークの構成が刻々と変化する。内部ネットワークは仮想スイッチによる柔軟な経路制御によって実現されるが、設定ミスなどによって意図しない経路が構築される場合がある (例えばファイアウォール用の VNF が迂回されるなど)。さらに、仮想スイッチは性能ボトルネックになる可能性が高く、仮想スイッチの負荷状態も重要な管理項目である。

NFV ノードでは、専用ハードウェア機器と同等の通信性能を実現するために DPDK を搭載する事が多い。効率化のため、CPU やメモリ、NIC などのリソースは DPDK が独自に管理するが、(i) 継続的な (受信) ポーリング処理のために CPU の使用率が常に 100%になる (ii) 巨大なメモリプールを作成し独自にメモリ割り当てを行う (iii) NIC の管理を排他的に行う (OS からは NIC が見えない)、という特徴がある。そのため、各種リソースの使用量に着目していた従来の監視技術では NFV ノードの状態把握ができず、通信障害を未然に防ぐための予防的な対策 (仮想機器のオートスケーリング等) の実施ができない。したがって、NFV ノード内部の特異なシステム構成を考慮しない従来のネットワーク/サーバ機器管理手法では、NFV ノードに起因する様々な通信障害の原因を究明できず、適切な対処ができない。本研究では、NFV ノード内の各コンポーネントを横断する共通基盤を開発し、内部構成の透明化によって一般的なネットワーク管理者による通信障害要因の特定を可能にする。さらに、DPDK のリソース使用方法に特化した状態把握の仕組みを確立し、NFV ノードにおける通信障害の予防に役立てる。

3. 研究の方法

本研究では、NFV ノードのシステム構成を考慮し、ノード内部におけるパケット転送経路や各過程の所要時間およびシステム負荷を可視化し、標準的なネットワーク管理者が NFV ノード内で発生する通信障害に対応できる仕組みを確立する。さらに、CPU などのリソース使用率に変わる新たな状態指標を考案し、NFV ノードの障害を予防する仕組みも確立する。本研究では、次の三つの課題の達成によって上記の研究目標を実現する。提案システムの全体像を図2に示す。

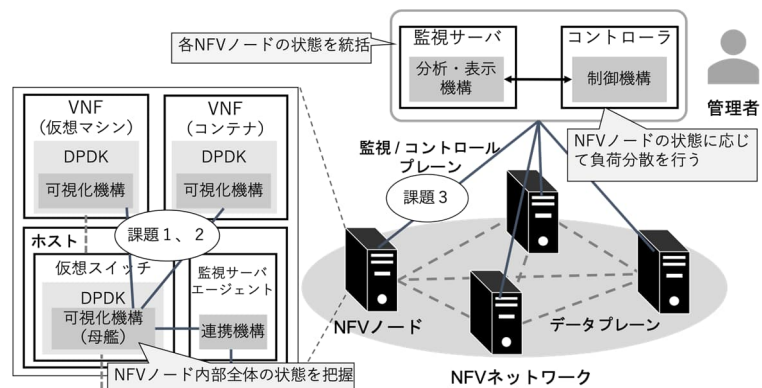


図 2 提案システムの全体像

課題 1: NFV ノード内部でのパケット処理過程や所要時間等を把握する仕組みの実現

提案システムを設計する際は、機能面の課題と実用面の課題の双方を考慮する必要がある。まず機能面については、メトリクス取得に伴う通信性能への影響を最小限に抑え、10 Gbps 超の通信トラフィックへの対応が必須である。また、コンテナ/仮想マシン形式の VNF (CNF) における内部状態を把握する術も求められる。続いて実用面については、保守運用や互換性のために、既存のシステムコンポーネントに対する変更を最小限に抑えたと共に、さまざまな VNF、仮想スイッチをサポートする必要がある。

DPDK は NFV における標準的なパケット I/O 機構であり、ホストマシンおよび VNF の双方において利用される。DPDK 自体は単なるライブラリ群であり、仮想スイッチプロセスや VNF プロセスはこれらを直接取り込む。そのため、本課題で実現する機能を DPDK 内部に実装する事で、既存のシステムコンポーネントに対する変更を最小化できる。DPDK はパケット処理のためにいくつかの CPU コアを占有するが、当該コア上で管理機能 (既存管理ツールとの連携等) を実行するとパケット処理性能が低下するため、管理機能専用の CPU コアを別途用意し、パケット処理用の CPU コアではアセンブリ命令による経過時間取得など、軽微な処理のみを実行する。コンテナ/仮想マシン内部の情報を取得する際は、Linux の標準機能である `vhost-vsock` を利用し、ホスト-仮想間の DPDK 同士でやり取りを行う。ホスト上の DPDK を母艦とする事で、NFV ノード全体の状態を把握する仕組みを実現する。

課題 2: DPDK のリソース使用方法や処理モデルに特化した NFV ノード状態指標の考案

DPDK の導入によって、`vmstat` などの既存ツールで得られる表面的な情報 (CPU 使用率やメモリ使用量等) はノード内部の状態把握には役立たなくなる。そのため、DPDK が内部でどの様にリソースを使用しているか、どの様な仕組みでパケット処理を行っているかを解析し、ノードの負荷と相関のある新たな指標を探り出す。

CPU 使用率等に変わる状態指標として、NFV ノード内部の受信キューにおける待機パケット数や、DPDK のパケット処理モデルである Rx-Process-Tx ループに要する時間 (1 回のループに要する CPU サイクル数) などが候補として挙げられる。そこで、様々な通信トラフィックを用いてそれぞれの指標値とノードの負荷状況の相関を調査し、障害予防という観点から最善の指標を決定する。

課題 3: 既存のネットワーク/サーバ管理ツール・プロトコルと連携する仕組みの実現

既にネットワーク/サーバ機器管理に関する膨大な成果物が存在するため、課題 1、2 で取得した情報は、既存の枠組みに則った形でネットワーク管理者に提供する必要がある。

既存のネットワーク/サーバ管理ツールでは、個々の機器にエージェントと呼ばれるプログラムを配置し、コマンドや OS のシステムコールを利用して必要な情報を取得する。そこで、DPDK 内部に提案手法専用のエージェント機能を実装し、既存管理ツール用のエージェントとやり取りするための標準 API を定める。その結果、既存管理ツールの枠組みの中で提案手法との連携が可能になる。また、SNMP には拡張 MIB という仕組みがあり、ネットワーク機器内の任意のコンポーネントに関する情報を表現可能である。本課題ではこのような仕組みを利用し、NFV ノード内部の仮想スイッチや各 VNF に関する情報を SNMP や Telemetry などの既存プロトコルによって取得する機構を実現する。

4. 研究成果

目標通り NFV ノード内部の詳細情報を取得するための状態指標を考案し、NFV-VIPP と呼ぶプロトタイプシステム (図 3) を開発してその効果を実証できた。具体的には、DPDK の Rx-Process-Tx ループ (パケット処理ループ) における消費サイクル数を監視し、過負荷状態の予兆を正確に捉えることができた (図 4)。NFV-VIPP システムのプロトタイプでは、 実際的なパケット処理を行うデータプレーンと 監視プレーンを分離する設計の採用によって、10 Gbps 級の通信トラフィックにも対応できるようになった。また、Zabbix などの既存の監視ツールと連携する機能も備えており、一般的なネットワーク管理者による NFV ノードの内部状態を考慮した運用管理を可能にするシステム構成になっている。

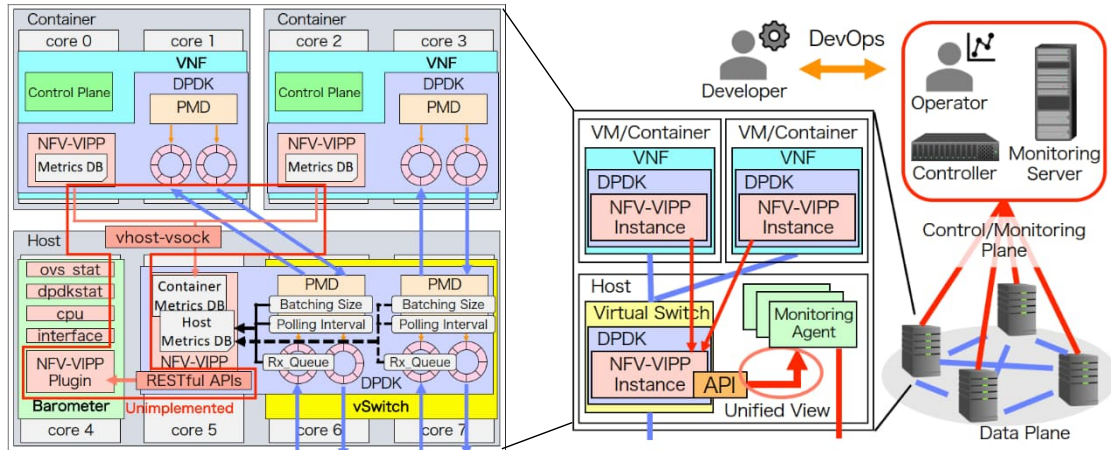


図 3 NFV-VIPP のシステムアーキテクチャ

NFV-VIPP システムの開発を経て、NFV システム内部の挙動について相当程度の知見が得られた。そこで、当初の研究課題をさらに発展させ、(汎用サーバを基盤とした) NFV の実用化に向けた障害となっている、パケット処理性能の問題についても取り組んだ。現行の環境において、コンテナ/仮想マシンとホストの接続点である仮想パケット I/O 機構が性能ボトルネックであると判明しており、厳密な解析の結果、パケット処理/転送に伴う CPU キャッシュの利用効率が性能向上の鍵を握っていると分かった。具体的には、ホスト環境と仮想環境をつなぐ vhost-user と呼ばれる仕組みに着目し、そのエッセンスのみを抽出した実験用プログラムを開発して様々な観点からの評価を実施した。実験用プログラムは受信プロセス、(単純なパケット転送のみを行う) NF プロセス、送信用プロセスの三つから構成される。解析の結果、まずキャッシュメモリの効率が最も性能に影響するのは NF プロセスであると判明した。そのため、送受信プロセスにおけるキャッシュメモリの利用効率は実性能には直結せず、vhost-user インタフェースの根本的な見直しが必要であるとわかった。

最後に、これまでの研究成果を踏まえ、(運用管理に適した) クラウドネイティブな性質を持つ NF (CNF) の実用化に向けた提言を行った。具体的には、ホスト (ベアメタル) 環境下での基本性能向上、仮想 (コンテナ) 環境下での基本性能向上に焦点を当て、CPU キャッシュと並列化を中核とした技術展望を示した。

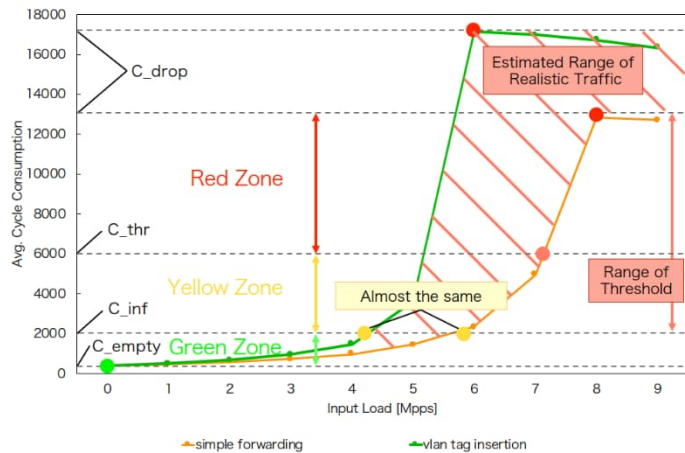


図 4 CPU サイクル数に基づいたシステム負荷の推定

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計5件（うち招待講演 1件 / うち国際学会 2件）

1. 発表者名 Masahiro Dodare, Yuki Taguchi, Ryota Kawashima, Hiroki Nakayama, Tsunemasa Hayashi, and Hiroshi Matsuo
2. 発表標題 NFV-VIPP: Catching Internal Figures of Packet Processing for Accelerating Development and Operations of NFV-nodes
3. 学会等名 2019 15th International Conference on Network and Service Management (CNSM) (国際学会)
4. 発表年 2019年

1. 発表者名 堂垂 正裕、川島 龍太、松尾 啓志
2. 発表標題 〔ポスター講演〕開発・運用サポートを目的としたNFVノード可視化機構の開発
3. 学会等名 電子情報通信学会 ネットワークシステム研究会
4. 発表年 2019年

1. 発表者名 Ryota Kawashima
2. 発表標題 A Vision to Software-Centric Cloud Native Network Functions: Achievements and Challenges
3. 学会等名 IEEE 22nd International Conference on High-Performance Switching and Routing (HPSR 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 竹谷 大地、川島 龍太、中山 裕貴、林 經正、松尾 啓志
2. 発表標題 仮想I/Oのキャッシュ汚染解消に向けたNon-Temporal命令の有効性評価
3. 学会等名 電子情報通信学会 情報通信マネジメント研究会
4. 発表年 2021年

1. 発表者名 川島 龍太
2. 発表標題 ソフトウェア指向による高性能クラウドネットワーキングへの展望
3. 学会等名 電子情報通信学会 ソサイエティ大会（招待講演）
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------