

クラウド環境における 自律運用制御基盤技術の研究開発

伝宝 浩史・加美 伸治
竹村 俊徳・柳沢 満

要 旨

クラウドコンピューティング環境が広まるにつれ、IT/ネットワーク資源運用の重要性が高まっています。例えば、構成変更の迅速性や資源配備の柔軟性が求められています。そこで、本研究開発では、新たな自律運用制御情報ネットワークアーキテクチャ、経路情報の共有手法、そして仮想化機器のプロファイリング技術の研究を進めてきました。これらを相互連動させていくことにより、クラウドコンピューティング環境における自律運用制御基盤の実現を目指します。

キーワード

●自律運用 ●自律制御 ●Publish/Subscribe ●情報共有 ●経路制御 ●プロファイリング

1. はじめに

クラウドコンピューティング環境が広まるにつれ、広域ネットワークを通して、計算資源をオンデマンドで利用でき、その利用者は、必要ときに必要な量のIT/ネットワーク資源を確保し、解放することができるようになってきました。本稿では、このような柔軟な資源運用が行われる環境でのIT/ネットワーク運用制御技術の研究開発について紹介します。

2. クラウド環境における自律運用制御基盤技術の研究開発

2.1 研究開発プロジェクト全体像

クラウドコンピューティング環境では、利用者のニーズに合わせて、資源量や構成を変更することが求められています。このような要求を満足するためには、観測・分析・対応・制御といった個別タスクの自動化に加えて、各タスクを相互連携させ、IT/ネットワークシステム自らが状況に合わせて、資源量や構成を迅速に最適化するための自律運用制御技術を構築していく必要があります。

この自律運用制御技術を実現するために、我々は3つの研究課題に取り組みました。1つ目の課題では、情報転送遅延の削減に着目しました。観測→分析→対応→制御という連携を迅速に行うためには、情報転送遅延の削減が求められます。そこで、OpenFlow技術を活用した自律運用制御情報ネットワー

クアーキテクチャとその情報転送機構を研究開発しました。

2つ目の課題では、ネットワーク経路制御における経路情報の共有手法に取り組みました。最適な経路を迅速に設定するためには、経路制御精度を維持しつつ、負荷を軽減することが必要になります。そのために、経路制御精度に影響を与えない範囲で、取り扱う情報量を削減する自律運用制御情報共有プラットフォームを研究開発しました。

3つ目の課題では、ネットワーク制御サーバの安定性向上に取り組みました。ネットワーク制御サーバを安定して稼働させるためには、負荷変動を吸収できる性能を持った仮想機器を選択する必要があります。そこで仮想機器の性能推定を行い、最適な仮想機器を選択するプロファイリング技術の研究開発しました。

以上のような技術課題の解決を通して、クラウド環境における自律運用制御基盤の実現を目指します。

2.2 自律運用制御情報ネットワークアーキテクチャ

この研究課題では、OpenFlow技術を活用した新たな自律運用制御情報ネットワークアーキテクチャとその情報転送機構を研究対象としています。前述したように、IT/ネットワーク資源の自律的な運用制御は、観測→分析→対応→制御というサイクルを繰り返します。この過程の中で、個々のタスクにおける情報の流れに着目してみると、**図1**に示すように、各々のタスクに応じて収集（多対1）型、検索（メッシュ）型、

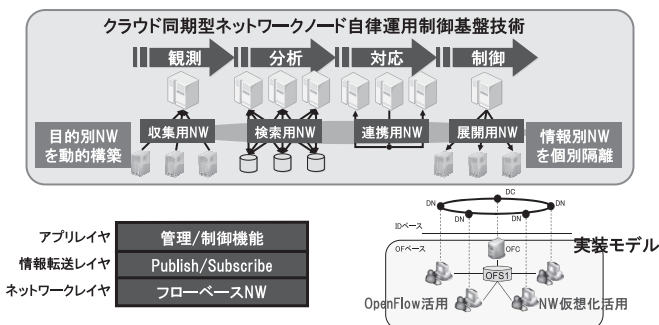


図1 自律運用制御情報ネットワークアーキテクチャ

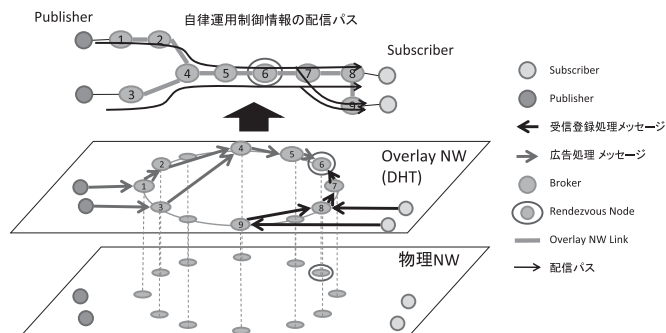


図2 分散イベントルーティング

連携（バス）型、展開（マルチキャスト）型などが考えられます。また管理内容に着目すれば、CPU、メモリ、ディスク、帯域、遅延など、多様な項目があり、管理目的に応じて、求められる情報転送品質も異なります。そこで、自律運用制御情報ネットワークアーキテクチャの要求条件として、管理目的別ネットワークの構築機能や管理目的別ネットワーク単位の資源割り当て機能が挙げられます。そこで、我々は、OpenFlow技術を活用し、管理目的単位にフローを定義し、動的にフローごとのネットワークを構築、更に個々の管理目的別ネットワークを隔離し、このネットワーク単位で資源割り当てを行う方法を考案しました。

状況に応じて柔軟に資源運用が行われる環境では、管理対象機器が増減するため、例えばネットワーク状態を観測する機能や、経路制御を行う機能においてオーバーヘッドが生まれます。そこで、このようなオーバーヘッドを削減するために、自律運用制御情報ネットワークとして、Publish/Subscribe通信モデルの適用を検討しました。この通信モデルは、Publisherというデータ送信機能、Subscriberというデータ受信機能、そしてこれらを疎結合する情報転送機能で構成されます。この疎結合通信では、PublisherとSubscriberは隔離され、情報転送機能を仲介役として通信を行います。これにより、クラウド環境における、構成の変化を運用制御サーバに隠蔽することができます。例えば、管理対象機器が増減したとしても収集情報が同じであれば、その増減はPublish/Subscribe層で隠蔽されます。このように資源の増減を許容しつつ、その特性をPublish/Subscribe通信層が吸収することで、シンプルなアーキテクチャとなります。そこで、その自律運用制御情報ネットワークノードとOpenFlowスイッチを制御するためのノードの機能設計を行いました。

Publish/Subscribe通信層の中では、管理情報の名前（情報名）に基づくルーティング方式の検討を行いました。本方式では、Publisherから投入された自律運用制御情報が、図2に示されるように、オーバーレイネットワーク上に構築された配信パスを通して、Subscriberへ転送されます。この配信パスはPublish/Subscribe通信における広告処理、及び受信登録処理によって、事前に構築されます。広告処理では、Publisherが情報をハッシュすることで、ユニークなIDを生成します。一方、受信登録処理では、Subscriberが受信したい情報名をハッシュしてIDを生成します。広告/受信登録処理で生成されたIDと同じ論理ノードIDを持つノードが、広告/受信登録先となるため、このノードが仲介者となってPublisher～Subscriberへ自律運用制御情報が配信されます。

以上のように、本研究課題では、OpenFlow技術を活用した新たな自律運用制御情報ネットワークアーキテクチャとその情報転送機構の実現を通して、情報転送遅延の削減、システム自律運用の向上に貢献することを目指します。

2.3 自律運用制御情報共有プラットフォーム

本課題では、経路制御精度に影響を与えない範囲で、取り扱う情報量を削減できる自律運用制御情報共有プラットフォームについて研究しました。例えば、あるネットワーク内のノード障害やリンクの過負荷状態が検出されたとき、そのネットワークは、周辺のネットワークと連携して、経路を変更することがあります。ここで、個々のネットワークを管理するサーバが、互いにすべての障害情報や負荷情報を共有すると、経路制御の精度は高くなりますが、その情報共有に伴う処理コストが大きくなります。そのため即応性やスケー

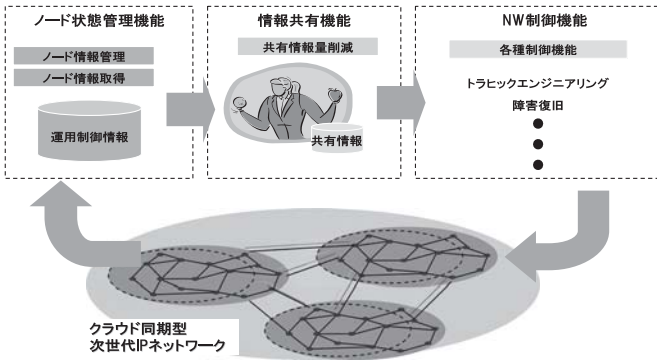


図3 ネットワーク運用制御ダイアグラム

ラビリティに悪影響が出ます。そこで、経路制御の精度を維持しつつ、経路制御に必要な情報を効率良く共有する仕組みが必要になります。

今回は、図3に示すように、個々のネットワークノードのリンク情報を収集するノード状態管理機能と、経路計算を行う経路制御機能という単純なモデルを用いて、両機能間で効率良くリンク情報を共有する情報共有機能の検討と、シミュレータ開発を行いました。

ネットワークシミュレータ (NS-2) でのシミュレーション検証では、リンク情報から経路制御を行う制御手法として QOSPF (要求帯域を満たす最小ホップ経路を選択し、その経路に沿って帯域保証を行う経路制御) を採用し、その情報共有部分に改修を加えて共有情報量をコントロールすることで、情報量削減のデータフローの品質に与える影響を評価しました。この改修は、あらかじめ決められた制御ポリシーに従って経路計算を行うようにしたものです。そしてネットワークモデルとして、5つのドメイン、そして各ドメインは10個のネットワークノードを有したモデルを用いて実行しました。その結果として、情報共有量 (1:100%共有、0:0%共有) に対する、各データフローのスループットの合計 (図4)、全リンク帯域のうち実際に利用している帯域の割合 (図5) を示します。これらは、30回の試行結果の平均値が用いられています。図4からは、情報共有量が60~70%程度まではスループット低下は少ないですが、それ以下ではスループット低下が大きくなります。つまり30%程度までであれば共有するリンク情報を削減したとしてもネットワーク品質に大きな影響がないことが分かりました。図5も同様で、70%を境に品質劣化が見られました。これは、70%を下回ると効率的な経路制御が

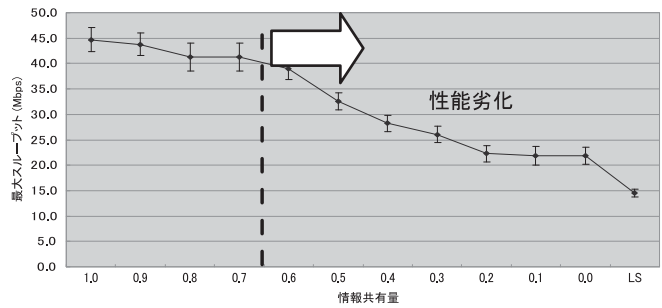


図4 情報共有量による最大スループットの変化

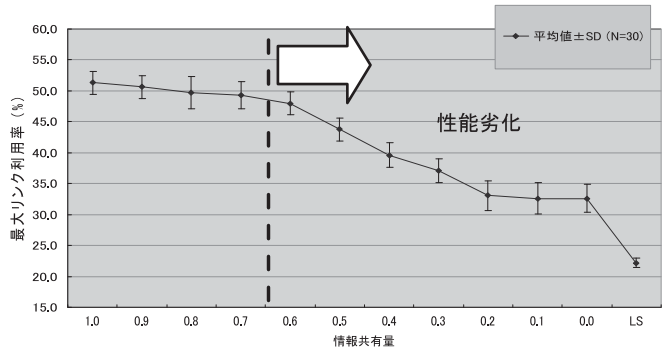


図5 情報共有量による最大リンク利用率の変化

できず、トラヒックが限られたリンクに集中することを意味しています。以上の結果から、ノード状態管理機能と経路制御機能間で共有すべきリンク情報量を100としたとき、30%程度のリンク情報削減であれば、ネットワーク品質に悪影響を与えないということが分かりました。

最後に、経路決定に有用な情報を生成するリンク順にIDを振り、ランダムに個々のリンクで観測されるリンク情報を欠落させた場合の最大スループットを測定しました。図6は、その測定結果であり、リンクID順に、測定された最大スループットを示しています。左側に示されたリンクに関するリンク情報が削除された場合、最大スループットが小さくなる傾向が出ています。そこで、左端から10個のリンク情報を削除すると、最大スループットは8.5Mbpsにまで減少しました。また逆に、右側のリンク情報が削除された場合、最大スループットが大きくなる傾向が出ています。そこで、右端から10個のリンク情報を削除すると、最大スループットは10.2Mbpsにまで増加しました。つまり、ランダムに30%のリンク情報を

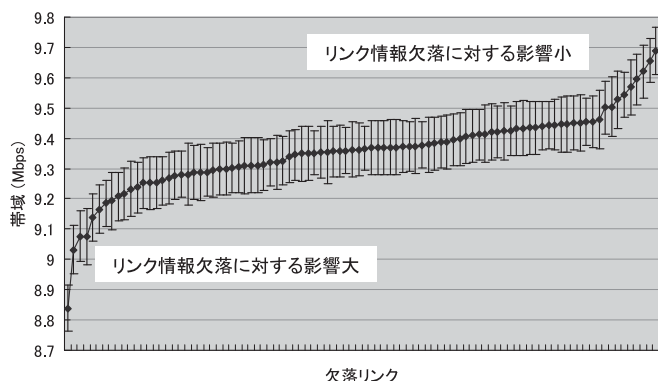


図6 リンク情報の欠落による総スループットへの影響

削減するよりも、厳選して30%のリンク情報を削減する方が、約20%の性能向上を見込めることが分かりました。

以上のように、これまでノード状態管理機能と経路制御機能における情報量削減に着目し研究開発を進めてきました。一般に情報量の増加に対して処理負荷は非線形に増加することが予想されるため、上記の共有情報量の削減は、数千から数万ノードレベルの大規模ネットワークでは実際の処理時間の削減に対し更なる効果が期待できると考えています。今後は、更に多くの運用制御項目に対して、本方式を適用するべく検討を進めていく予定です。

2.4 仮想リソースプロファイリング技術

本研究課題では、自律運用制御基盤の安定性を向上するための研究開発を行っています。一般的に、安定したシステムは、稼働率が高く、高い可用性を持っています。そのようなシステムは、個々でエラーや故障が発生しても、全体としてシステムダウンに至らないようクラスタリングなどの技術が適用されています。しかし仮想化技術が適用されたクラウド環境下では、新たな工夫が必要となります。クラウド環境で仮想化されたIT/ネットワーク機器が利用する物理資源は、異なった性能、異なった容量、そして異なった位置に設置された物理機器やデバイスから提供されることもあります。このような状況で、仮想機器を安定運用していくためには、その仮想機器の資源消費特性を用意し、より適した物理資源の候補を選定、その候補の中から最適な物理資源を選択する必要があります。

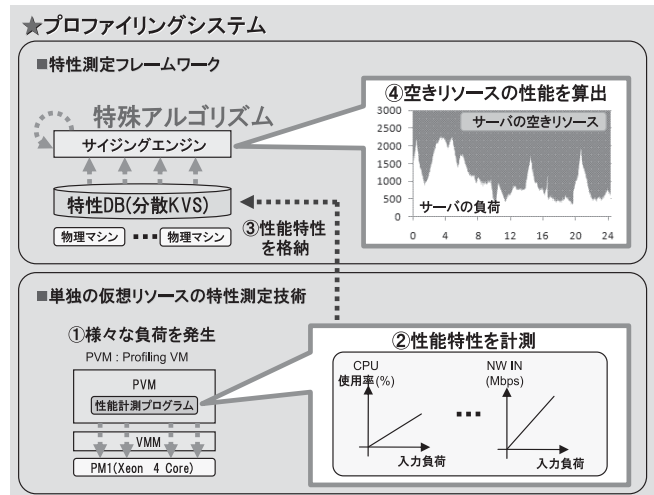


図7 仮想リソースプロファイリング技術

そこで本研究開発では、図7に示すように、プロファイリングのためのさまざまな負荷を発生させ、ネットワーク制御サーバを稼働する仮想機器の性能特性を計測する「特性測定技術」と、計測された性能特性データと物理機器の稼働状況から、クラウド環境内に散在する空きリソースの処理性能を推定・比較する「特性測定フレームワーク」で構成される仮想リソースプロファイリング技術を開発しました。

特性測定機能における性能特性の分析方式としては、2種類のプロファイリング方式を検討しました。第1の方式は、CPU負荷やネットワーク送受信などのリソース単位に擬似負荷をかけ、その実測値から性能特性の近似関数を算出して細かい粒度の性能特性値を生成し、その値の中から仮想リソースの処理性能を推定します。第二の方式は、アプリケーションの処理によく利用される複数のシステムコールを同時に実行することによって、システムコール間の影響を含めた性能特性を実測し、そこから仮想リソースの処理性能を推定します。

これらの技術の効果を確認するため、2台の物理サーバSV#1、SV#2に対して性能特性データの収集と負荷指標の推定を行うシステムを試作し、I/O仮想化によるCPU負荷を想定した2種類の負荷（CPU、ネットワーク出力）による評価を行いました。以下に、第一の方式の計測結果を示します。

図8は、SV#1への入力負荷として仮想マシン上でWebサーバを稼働させたときのSV#1のCPU負荷を示し、図9は、SV#2上で同等の処理を行った場合のCPU負荷の実測値と、2

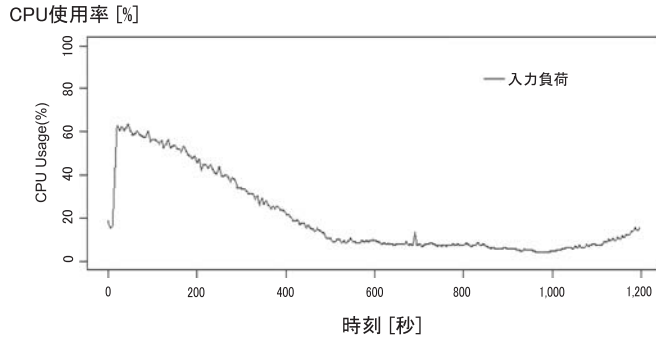


図8 SV#1での入力負荷

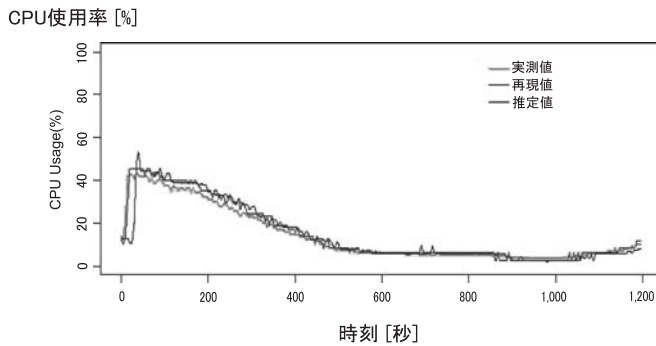


図9 SV#2での推定結果

台のサーバのプロファイリング結果からSV#2上での負荷を算出した推定値と、SV#1のプロファイリング結果から入力負荷に相当するSV#1上での処理量を推定し、その処理をSV#2上で実行して負荷を擬似的に再現させた再現値の3つを示しています。図8、図9から、SV#1とSV#2では、CPU性能に1.5倍程度の性能差があるため、処理能力の高いSV#2が低いCPU使用率を示す（例えば、SV#1上で約60%のピーク負荷がSV#2上では約40%となる）ことを確認しました。図9から、本方式の推定値と再現値として、実測値に近い値を見積もることができ、性能差のある物理サーバ間でも負荷が推定できています。そして、推定値と再現値がそれぞれ実測値に対してどの程度離れているかを表す乖離率は2%程度であり、十分に負荷の推定値として利用できることが確認できました。

3. おわりに

本稿では、クラウドコンピューティングに適したIT/ネットワーク運用制御を構成する3つの要素技術として、自律運用制御情報ネットワーク、自律運用制御情報共有プラットフォーム、仮想リソースプロファイリング技術について報告しました。今後、これらの技術の実用化を目指し、NECのクラウドコンピューティング製品の強化に貢献します。

なお本研究は、総務省の委託研究「セキュアクラウドネットワークワーキング技術の研究開発」プロジェクトの成果です。

執筆者プロフィール

伝宝 浩史
システムプラットフォーム研究所
主任

加美 伸治
システムプラットフォーム研究所
主任

竹村 俊徳
サービスプラットフォーム研究所
主任研究員

柳沢 満
サービスプラットフォーム研究所
主任研究員