

特集号
招待論文

サービス品質保証型の 大規模省電力クラウドシステム

—3都県に跨る大規模システムによる実証実験—

鈴木 敏明^{†1} 垂井 俊明^{†1} 馬場 智宏^{†1} 高瀬 晶彦^{†1} 山田 亜紀子^{†2}
清水 翔^{†2} 宗宮 利夫^{†2} 竹下 秀俊^{†3} 堀江 光^{†3} 山中 直明^{†3}

^{†1}(株) 日立製作所 ^{†2}富士通 (株) ^{†3}慶應義塾大学

近年、サーバ、ストレージ、端末、ネットワーク等の技術進展により、クラウドサービスが急激に普及し始めている。また、サービスの多様化やユーザの増加により、サービス提供システムの拡大とともに消費電力が急増しており、システムの省電力化が重要な課題の1つとなっている。本論文では、複数のデータセンタとそれらを接続するネットワークより構成される大規模なクラウドシステムおよび提供するサービスに対して、サービスの品質を維持したシステムの省電力化手法について提案する。また、3都県に跨る大規模システムによる実証実験により、サービス品質を維持したクラウドシステムの省電力化について有効性を検証した。

1. はじめに

近年、サーバ、ストレージ、ユーザ端末等の性能向上やネットワークの伝送帯域拡大により、インターネット等のネットワークを介して遠隔にあるデータセンタ(DC)より提供されるクラウドサービスの利用が急拡大している。一方、クラウドサービスの普及とともに、サービスを提供するDCの建設が増加し、DCによる消費電力の拡大が予想[1]されている。また、サービスを遠隔地から提供するためのネットワークによる消費電力の急増も予想[2]されており、DCやネットワークによる消費電力の削減が重要な課題となっている。

本論文では、複数のDCをネットワークで接続した大規模なクラウドシステムおよび提供するサービスに対して、提供するサービスの品質を維持したシステムの省電力化手法について提案する。また、3都県に跨る大規模なシステムによる実証実験により、提案システムの有効性を検証する。

以降の章構成は、以下の通りである。第2章において、従来のクラウドシステムにおける省電力化手法の課題について述べる。第3章では、クラウドシステムを省電力化するための要求条件を規定した後、ネットワークとDC連携による省電力制御について提案し、また実証実験および評価について述べる。第4章では、サービスレイヤによる省電力制御について提案し、また実証実験および評価について述べる。最後に第5章にて結論を述べる。

2. クラウドシステム省電力化の課題

クラウドシステムは、一般にアプリケーションサービス等を提供するサーバやストレージ等から構成されるDC、およびDC間やユーザとDC間を接続するネットワークから構成される。クラウドシステムを省電力化する場合、ネットワークによる消費電力を削減する方法とDC内のサーバ等による消費電力を削減する方法とが考えられる。

従来の消費電力削減方法[3],[4],[5]では、上記のサーバによる消費電力の削減とネットワークによる消費電力の削減が主に独立して研究開発されている。しかし、サーバとネットワークによる消費電力削減を個別に制御した場合、提供するサービスの品質が大きく劣化するという課題が発生する可能性がある。

たとえば、DC内においてサービスを提供する仮想マシン(VM: Virtual Machine)の負荷に着目し、VMの負荷が低下した状況においてサーバ上で稼働させるVMを特定のサーバに片寄せし、不要なサーバの電源を停止する等による省電力化方法が研究開発されている。本方式の場合、物理サーバが提供可能なCPUの負荷のみに着目した片寄せによって、多数のVMが同一のサーバ上へ集約されることで、VMの移動に伴いネットワークを介して伝送されるデータ量も集約されるため、ネットワーク上に輻輳が発生し、提供するサービスの品質が劣化するという課題が発生する。

ムは、複数のDCをアクセスネットワークで接続し、さらに複数のアクセスネットワークをメトロネットワークで接続した大規模な階層構造を有している。提案するクラウドシステム内の主なサーバの役割は、下記の通りである。

- **メトロネットワーク管理サーバ**：メトロネットワークのトポロジを管理し、流出入するトラフィック量に応じて転送経路の適正化を図り、メトロネットワーク内における省電力化を図る。また、アクセスネットワーク間でVMの移動が発生した場合に、VMが移動した先のアクセスネットワークへデータを伝送するための制御を行う。
- **アクセスネットワーク管理サーバ**：アクセスネットワークのトポロジを管理し、また流出入するトラフィック量からアクセスネットワークの各経路が消費する帯域を予測する。また、伝送量に基づき経路の集約を図り、アクセスネットワーク内における省電力化を図る。さらに、DC間にてVMの移動が発生する場合に、VMが移動した先のDCへデータを伝送するための制御を行う。
- **ネットワークモニタリングサーバ**：アクセスネットワークにおけるトラフィックの流出入量を監視し、各リンクにおける帯域消費の状態をモニタリングし管理を行う。
- **DC統合管理サーバ**：複数のDCを統合的に管理し、深夜や早朝といったVMの負荷が低下する時間帯において、VMのサーバ間での再配置をシステム全体で実行するようDC管理サーバへ通知し、省電力化の実行開始、終了を制御する。また、DC管理サーバからの要求をアクセスネットワーク管理サーバへ通知し、かつアクセスネットワーク管理サーバ

からの応答をDC管理サーバへ通知する。

- **DC管理サーバ**：DC内に存在するサーバ、VM、スイッチにおける処理負荷量やトラフィック量をモニタリングし、サーバにおける将来の処理負荷量やスイッチにおける消費帯域を予測する。また、他のDC管理サーバと連携し、負荷の低下したVMの移動先候補となるサーバを、将来のサーバ負荷とネットワーク側の将来負荷予測とから選定し、サービス品質の維持と省電力化を両立したVMの配置を決定する。

3.3 省電力制御

3.3.1 アクセスネットワークとデータセンタ連携による省電力制御

本項では、サービス品質を維持したクラウドシステムを省電力化するための、アクセスネットワーク管理サーバとDC管理サーバとが連携したVM配置決定方式について述べる。具体的には、DC内のVM負荷が低下した時間帯に、同一のサーバ上で稼働するVM数を増加させ、不要なサーバを停止することによりDCにおける省電力化を実行するが、VMから提供するサービスの品質を劣化することなく省電力化を達成するVMの配置算出方式について述べる。

図2に、サービス品質を劣化することなくクラウドシステムの省電力化を達成する、VMの適正配置算出プロセスを示す。

- ① DC統合管理サーバからDC管理サーバへ、VMの処理負荷が低下する時間前に、順次VMの配置適正化指示を送信する。DC統合管理サーバでは、VLANごとにVMが移動可能なDCリストを保持しており、DC管理サーバへVM配置適正化指示を送信する際、移動可能DCリストも併せて通知する。
- ② DC管理サーバは、自DC内で稼働しているVMごとの配置適正化順序を決定する。たとえば、VLANごとやCPU単位当たりの消費電力量の多い順に従って決定する。
- ③ DC管理サーバは、自DC内の移動を予定しているVMが、将来に必要とするリソース量（CPU処理量、メモリ量、送受信帯域）を提供可能な他DC内のサーバ候補を、並列処理にて他の複数のDC管理サーバから非同期に受信する。
- ④ DC管理サーバは、他のDC管理サーバより受信したすべてのVMの移動先候補となるサーバに対して、VMを移動した場合に、アクセスネットワーク

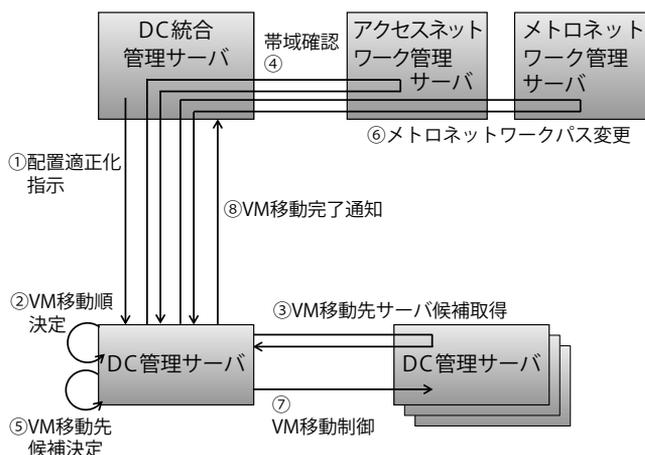


図2 サービス品質を維持した省電力 VM 配置算出プロセス

としてVMが必要とするネットワーク帯域を提供可能かをアクセスネットワーク管理サーバへ確認する。

- ⑤DC管理サーバは、移動を予定しているVMが将来に必要なとするリソース量と、移動先候補となっているサーバにおける将来の処理負荷量およびアクセスネットワークにおける将来の空き帯域量を考慮し、リソースの不足が発生しないように、実際のVM移動先候補となるサーバを決定する。
- ⑥DC管理サーバは、VM移動先候補のサーバを決定した段階において、アクセスネットワーク間でVM移動が必要な場合は、DC統合管理サーバおよびアクセスネットワーク管理サーバを経由し、メトロネットワーク管理サーバに対して、データ伝送パスの変更を要求する。具体的には、移動後のVMにユーザがアクセスできるように、かつ移動前と同等のアクセス帯域を確保するようにメトロネットワーク管理サーバへ要求する。メトロネットワーク管理サーバは、要求されたパスを提供可能かを判断し、DC管理サーバへ回答を通知する。
- ⑦DC管理サーバは、すべての移動を予定しているVMに対して、移動可能なサーバを判断した後、実際のVM移動を開始する。
- ⑧DC管理サーバは、自DC内の移動を予定しているVMすべての移動が完了した後、DC統合管理サーバに対して自DCの省電力化制御が完了したことを通知する。

3.3.2 メトロネットワーク省電力制御

大規模クラウドシステムでは、広域に分散配備された複数DCをアクセスネットワークおよびメトロネットワークによって接続する。本項では、大規模クラウドシステムの省電力化を実現するための、メトロネットワーク省電力化について述べる。

(1) メトロネットワーク制御アーキテクチャ

3.1.2項で挙げた要求条件を実現するための集中制御型アーキテクチャを図3に示す。メトロネットワーク管理サーバはメトロネットワーク全体の機器構成やリソースを管理しており、アクセスネットワーク管理サーバからのVM移動要求等のサービス要求シグナリングを受け付けるインターフェースを提供し、要求を満たす省電力なネットワークリソースを決定するアルゴリズムを備えたネットワークリソースコントローラ、および決定した経路での通信を実現するための機器コントローラを持つ。

(2) マルチレイヤ省電力経路決定アルゴリズム

メトロネットワークにおいては、電気的なパスと光パスのマルチレイヤネットワークで構成されるため、マルチレイヤを考慮した最適経路選択を行うことがネットワーク全体の省電力化を実現する上で有効となる。しかしマルチレイヤを考慮することで取り扱うべきパラメタが増大するため、計算時間が増大する懸念がある。そこで、現実的な時間で最適経路選択を実現するためにはヒューリスティックなアプローチが有効であり、すでにIP/WDMネットワークの省電力経路決定アルゴリズムを提案、評価している[6]。

既提案アルゴリズムを用いると、図4に示すように消費電力が多いルータをカットスルーする光パスをトラフィック量や距離を考慮して適切に設定し、メトロネットワークリソース要求に対するブロッキング率を低く抑えつつネットワーク全体の消費電力量を抑えることが可能となる。提案方式の特徴としては、距離によって必要性が決定する光再生中継器の使用可能性を表現し、波長利用時やルータ利用時の電力特性をノードおよび辺のコストとして与えた補助グラフを生成し、最小重み経路問題に帰着した点である。これによりマルチレイヤでの省電力経路決定をダイクストラ法を用いて高速に行うことが可能となる。

(3) メトロネットワーク管理サーバの制御機能

アクセスネットワーク管理サーバと連携するメトロネットワーク管理サーバの動作を以下に示す。

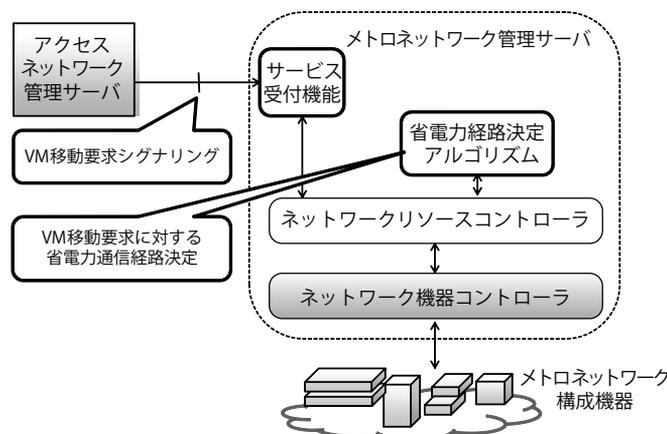


図3 メトロネットワーク制御アーキテクチャ

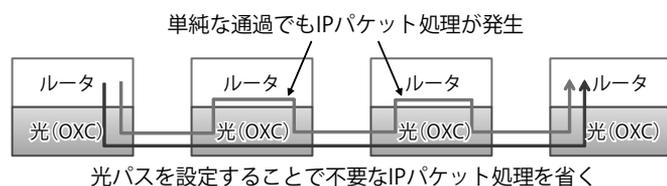


図4 光パスによるルータのカットスルー

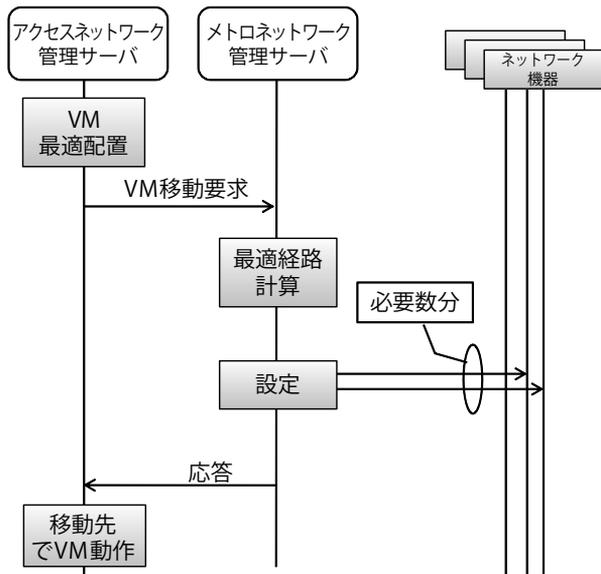


図5 VM 移動要求受信時の処理フロー

- ①メトロネットワーク内で通信を行うVMについての情報を取得し、メトロネットワーク初期通信設定を実施する。
- ②アクセスネットワーク管理サーバから、VM 移動要求シグナリングを受ける。
- ③要求内容からVM 移動先での通信区間や帯域を確定し、省電力経路を決定する。
- ④機器への設定を実施する。
- ⑤アクセスネットワーク管理サーバへ移動要求に対する処理が完了した旨応答を返す。

VM 移動要求を受けてからのフローは、図5に示した通りである。またVM 移動に伴う通信区間と帯域を確定するために必要な情報はアクセスネットワーク管理サーバから取得する。主な取得情報を表1に示す。

3.4 ネットワークとデータセンタ連携による省電力化評価

本節では、提案システムにおける広域での実証実験と結果について述べる。図6に実証実験の概要を示す。評価システムは、図に示すように宮城県、東京都、神奈川県に、それぞれ模範的に配備したDCをJGN-X^{☆1}

☆1 (独) 情報通信研究機構が推進する新世代ネットワークの研究開発を支えるテストベッド。

表1 アクセスネットワーク管理サーバから取得する主な情報

事前取得情報	異なるアクセスネットワーク同士の通信要件 (通信区間, 必要帯域) アクセスネットワークとVMの対応情報
移動要求時取得情報	移動VMのIPアドレス 移動VMの最大利用帯域 (双方向)

で接続した大規模なクラウドシステムである。

実証実験では、クラウドシステム全体において、サービス品質を維持した省電力化制御機能の検証を行う。また特に、1) メトロネットワーク、アクセスネットワークおよびDCが連携したサービス品質を維持した省電力化制御、2) 複数のアクセスネットワークを接続するメトロネットワークにおける省電力化制御について、機能の検証評価を行う。

3.4.1 実証実験システム構成

図7に実証実験システムの構成を示す。図に示すように、JGN-Xをメトロネットワークに見立てて複数のアクセスネットワークを接続している。メトロネットワークでは、各アクセスネットワークを接続するメトロネットワークのノード (光ノード, ルータ, エッジスイッチ) とそれらを管理制御するメトロネットワーク管理サーバから構成されている。また、アクセスネットワークは、メトロネットワークやDCと接続するスイッチとアクセスネットワーク管理サーバから構成される。DC1~4は、多数のVMを収容するサーバ群、DC管理サーバ、およ

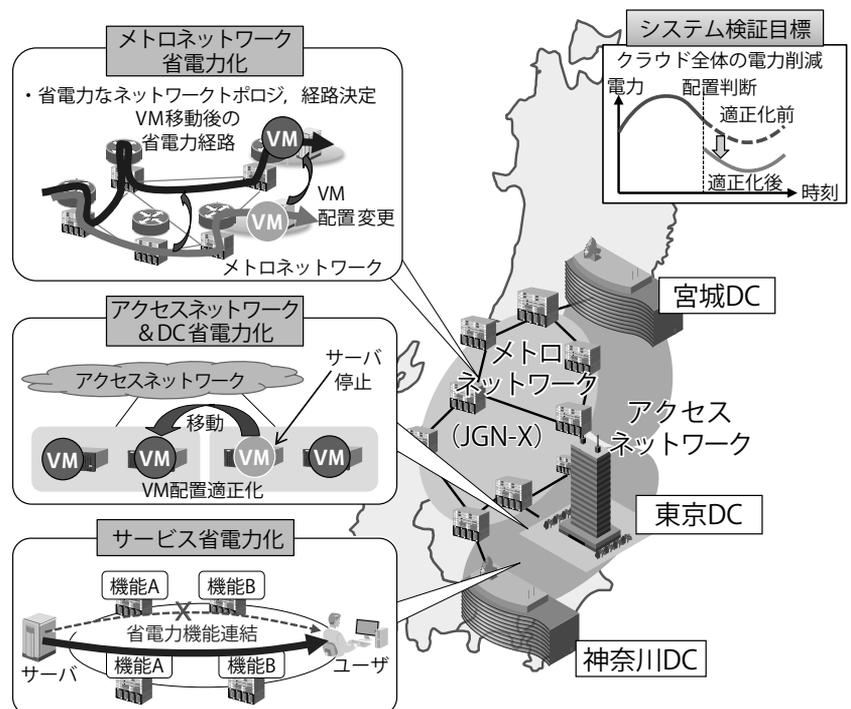


図6 実証実験概要

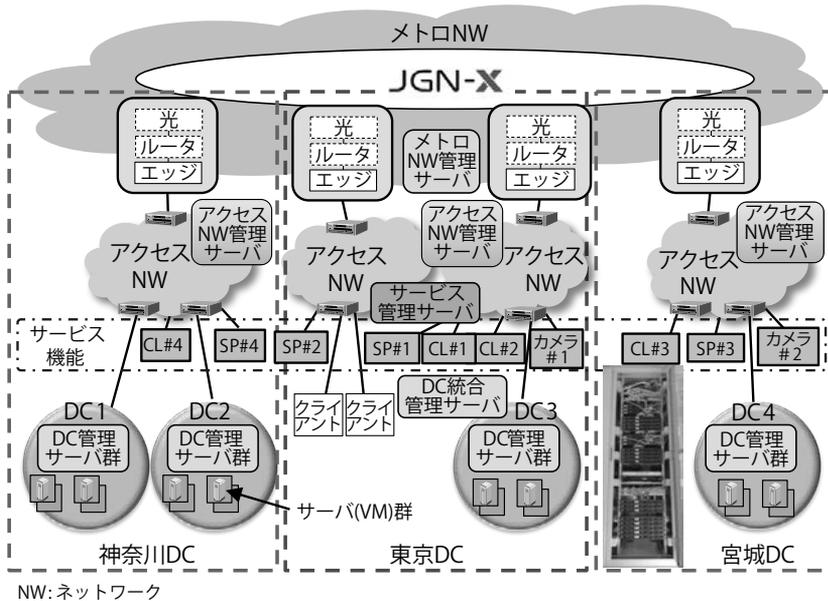


図7 実証実験システム構成

びユーザ管理サーバにより管理制御される。さらにDC全体は、DC統合管理サーバにより統合管理制御される構成である。

(1) アクセスネットワークおよびDCの実装

図7に示した実証実験システムを、実機およびエミュレータから構成した。具体的には、アクセスネットワークや各種の管理サーバについては、実機にて構成した。一方、DC内のスイッチやサービスを提供するサーバは、エミュレーションソフトウェア[17]により構成した。

(2) メトロネットワーク管理サーバ実装

本実証実験では、メトロネットワーク管理サーバをOpenFlow^{☆2}[7]による制御を実施するものとして実装した。ネットワーク機器コントローラとしてはBig Switch Networks社のOpenFlowコントローラFloodlight[8]を用い、ネットワークリソースコントローラをFloodlightアプリケーションモジュールとして実装した。制御対象をOpen vSwitch[9]としたため、光パスの設定は実際には行わず、想定光ネットワークトポロジに対しての経路計算を実施し、結果をメトロネットワーク管理画面に表示するのみとした。

3.4.2 省電力化制御の評価

(1) ネットワークとデータセンタ連携による省電力化評価

DC1～4内に合計で1,000VM（400サーバ）を模擬的に配備し、クラウドシステムとしての省電力制御機能を検証した。400台の物理サーバを実際に配備すること

は、本実証実験では困難であったため、エミュレーションソフトウェアにより物理的なサーバを模擬した。また検証実験では、3都県に跨るクラウドシステム構築や1,000VMに対する負荷モデルを準備する必要等、当初の想定以上の工数を必要とする大規模な実験となった。

実証実験結果としては、VMの配置適正化制御により、アクセスネットワーク間でVMの移動が必要と判断された場合に、メトロネットワーク管理サーバへデータ伝送パスの変更指示を出すことで、配置適正化後においてクライアントからVMへの通信が維持されることを確認した。また、VM配置適正化後においても、ネットワークのリンク帯域を超過するよ

うなVMの移動が行われていないことを確認した。アクセスネットワークにおける消費電力削減の検証では、伝送負荷の低減により、各アクセスネットワークに配備した9台のスイッチのうち、3台から4台が停止となり、全体として20.2%の消費電力削減が可能であることを検証した。また、VMの利用負荷が深夜や早朝に低下する処理負荷モデルを用いて4DCにおける省電力制御の効果を検証した結果、平日パターン（5日分）では、24.3%の省電力化を、休日パターン（2日分）では、46.6%の省電力化を検証しており、1週間の単位では、30.6%の消費電力を削減可能なことを検証した。結果として、システム全体で20%～30%の消費電力の削減が可能であることを検証した。

(2) メトロネットワークとアクセスネットワーク連携制御の評価

VMがメトロネットワーク越しに通信するためのメトロネットワーク初期設定処理性能を評価するため、DC1～4内に配備されるVMの数を1,000～10,000まで変化させたときのメトロネットワーク初期通信設定に要する時間を測定した。結果として4つのエッジノードからなる検証ネットワークでは10,000VMの初期通信設定を0.4秒程度で処理することを確認した。また設定に要する時間が支配的であったが、これは検証ネットワークが単純なトポロジであり経路計算時間が高速に完了していたためといえる。

またメトロネットワーク管理サーバのVM移動要求処理性能を評価するため、VMが移動する際の処理時間を測定した。具体的には、連続して100個のVM移動要求をアクセスネットワーク管理サーバから受信し処理

☆2 OpenFlowは、Open Networking Foundationの商標または登録商標です。

時間を確認した。結果として移動要求処理時間は平均 3.8msec であり、1 秒間に約 250 程度の移動要求を処理することが可能なことを検証した。

4. ネットワーク情報に基づくサービスレイヤによる省電力制御

第2章で述べたように、ネットワークとサーバの消費電力の総和が最少となる構成でサービスを省電力で提供する仕組みが必要となる。本章では、ネットワークのトポロジ情報（リンクの帯域、消費電力含む）を基に、サービスレイヤでの消費電力最小ルートの選択法、およびサーバの最適配置法について述べる。

4.1 要求条件

- サービス提供時、ネットワークとサーバを含めた消費電力が最少となるルートを選択する機能
- 省電力化と SLA (Service Level Agreement) を両立させる機能

4.2 サービス提供時の省電力制御

4.2.1 省電力マッシュアップサービスルーティング

マッシュアップサービスとは、ネットワークに接続したデバイス（サーバ内の処理プログラム、コンテンツ、ストレージ、カメラ、ディスプレイ、センサ等）をサービスパーツと呼び、サービスパーツを順次接続していくことによって新たなサービスパーツを創造することをいう。図8のマッシュアップサービス概念図に示すように、カメラで撮影された画像はサービスパーツBサーバに転送されて顔認識され、続いてサービスパーツCサーバに転送されて顔認識された部分が高解像度化され、ユーザには撮影画像が加工された画像（顔認識+高解像度化）が転送される。

マッシュアップサービスを最小消費電力で提供するには、サービスパーツとサービスパーツを接続するネットワークの消費電力の総和が最少となるようにルートを選択することが必要となる。ところが、両者を統一的に扱い消費電力の最小となる組合せを容易に選択することができない。一方、IPアドレスはIPv4のアドレス空間の枯渇に伴いIPv6化が進み、膨大なアドレス空間（約 2^{128} = 約 340 澗）を手に入れることができるようになった。

そこで、サービスパーツをノードとして扱い

IPアドレスを付与し、ネットワーク情報を基にサービスパーツをノードとして加え、消費電力をリンクコストに反映したサービスレイヤのトポロジを生成する手法を提案する。図9の上部は物理トポロジを示す図であり、レイヤ2スイッチ3個が各々リンクコスト（消費電力）2, 5, 7の3本のリンクで相互接続され、レイヤ2スイッチ#1にコスト5のリンクでコスト40の高解像度化ソフトを格納したサーバ#Aとコスト5のリンクでコスト10の受信サーバが接続され、レイヤ2スイッチ#2にコスト5のリンクでコスト30の顔認識ソフトが格納されたサーバ#Bと、レイヤ2スイッチ#3にコスト5のリンクでコスト15のカメラ用送信サーバが接続されている。図9下部は、サービスレイヤのトポロジを示す図であり、ネットワークレイヤのトポロジに、各サービスパーツがノードとして追加され、各サーバ内のサービスパーツのコストはリンクコストとして表現されている。

以上に述べたサービスレイヤのトポロジ生成により、ネットワーク層のトポロジと同様に、さまざまなルーティングプロトコルが活用できるようになり、最少消費電力の組合せを導出することが可能となる。

省電力のサービスルーティングアルゴリズムについては、線形計画法、拡張Dijkstra法（3D-Dijkstra法）[10]、さらにサービスパーツのコピーを考慮したアルゴリズム[11]、サービス開始時点でのサービス中のトラフィックの保留時間を考慮した動的経路選択手法[12]等が研究開発されてきており、本実証実験では省電力サービスルーティングに動的経路選択手法を用いた。

4.2.2 SLAを保証した省電力サーバ動的配置

省電力化とSLAの保証を両立させるには、クラウド間で負荷情報および電力情報を管理し、ユーザのリクエストを分散することが必要となる。

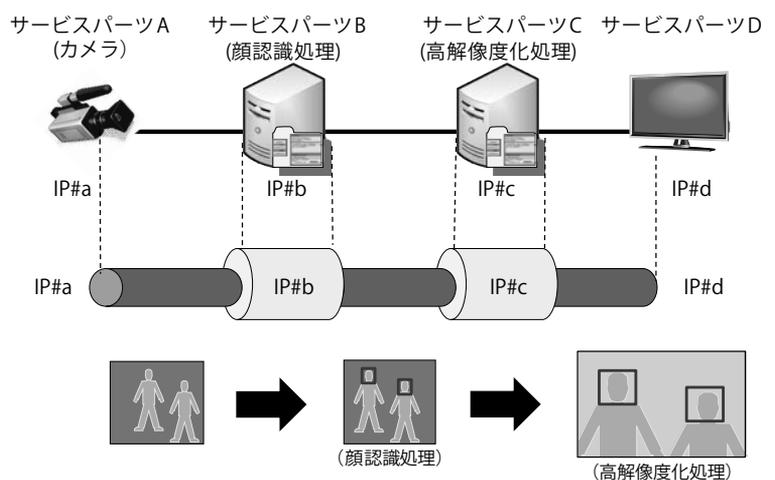


図8 マッシュアップサービス概念図

SLAを保証するには、1) 各クラウド上に適切な台数のサーバを配置すること、2) それらのサーバに対してリクエストを適切に分散させることが必要である。たとえば、サービスの可用性を保証するには、各クラウドはクライアントからのリクエスト量に応じた適切な台数のサーバの配備が必要となる。また、サービスの応答性を保証するには、要求された応答時間を満足すべく極力クライアントの近くにあるサーバでリクエストを処理する必要がある。

サーバ動的再配置基盤の開発には、既存のクラウド環境が一般的に提供するオートスケーリングの技術（クライアントからのリクエスト量に応じてサーバ数を自動的に増減する機構）を活用した。これにより、クラウド間連携を想定した場合、各クラウドへ割り振るリクエスト数を調整することでサーバを再配置することができる。

そこで提案手法では、SLAを考慮して各クライアントからのリクエストを適切なクラウドへ割り振る仕組みを導入する。具体的にはクライアントからのリクエストを適切なクラウドへ転送する役目を持つマッピングノードと呼ぶマシンを各クラウドに配置し、これらをオーバーレイネットワークで接続したマッピングノード群によってリクエストの割り振りを行う（図10）。マッピングノード間では保有する各サービスのSLAに関する情報を定期的に交換する。各クライアントがサービスにアクセスする際は、図10のように、まず任意のマッピングノードへリクエストを送り、マッピングノードがSLAを考慮して適切なクラウドへリクエストを転送することで、SLAを保証しつつサービスを提供する。

上記機構において、SLAを保証可能なリクエスト転送先が複数存在する際に、最も省電力効果が高いクラウドを選択することで、SLAを満たしつつ省電力効果を最大化する。省電力効果を考慮する際は、サーバを配置する物理マシンの電力量だけでなく、クライアントとの通信に要するネットワーク機器の消費電力も考慮し、漸近的にサーバの配置を改善していき最善の状態へと近づけていく[13],[14]。具体的には、SLAを満たすことができる配置をまず求め、その範囲内で省電力効果を高める配置へと徐々に変更していく形をとる。

提案手法においてマッピングノード群は分散ハッシュテーブル（DHT: Distributed Hash Table）を利用したオーバーレイネットワークを形成する。DHTは、キーとバリューのペアを

データとして持ち、それを複数のノードで分散管理する仕組みである。DHTでは、検索のためのインデックス情報などを管理するサーバが存在せず、各ノードが自律分散的に動作してサーバとクライアント両方の機能を果たしている。したがって、ユーザ数が増加しても、DHTを構成するノードの数を増やすことで対処可能であり、スケーラビリティが高い。上記機構においては、DHTの代表的なアルゴリズムの1つであるChord[15]を利用する。Chordは耐障害性や検索応答性等の点で優れたアルゴリズムであり、本手法との親和性が高い。

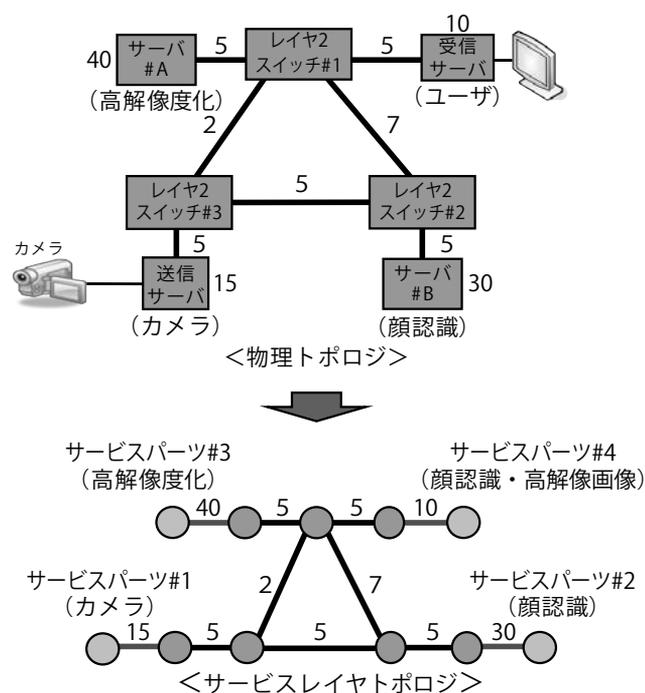


図9 サービスレイヤトポロジ生成

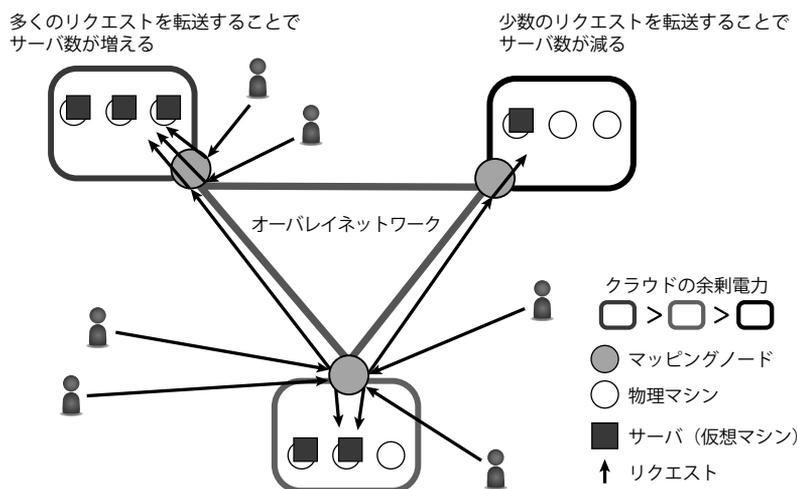


図10 リクエスト分散によるサーバ動的再配置

4.3 実証実験のシステム構成

本節では実証実験システムの構成について示す。図7に示すように、サービスレイヤでの省電力制御を検証するため、アクセスネットワークにサーバ設備を接続している。具体的には、アクセスネットワークにおいて、サービスマッシュアップ検証用に、複数のクライアント(CL#1～#4)、サービスパーツ(SP#1～#4)、カメラ(カメラ#1～#2)を接続している。

4.4 省電力化制御の評価

本節では、提案システムにおけるネットワーク情報に基づいたサービスレイヤによる省電力制御の実証実験と結果について述べる。

4.4.1 省電力マッシュアップサービスルーティング

図7の実証実験システム構成より、アクセスネットワークおよびメトロ/コアネットワークのトポロジと消費電力情報を基にして、サーバ内のサービスパーツ(アプリケーションソフト)をノードとして加えた図11に示すサービスレイヤトポロジを生成し、動的経路選択アルゴリズム[12]を用いて、最少消費電力となる組合せを導出し、マッシュアップサービスが提供できることを検証した。図11に示す例では、CA#1サーバとCL#2サーバ間は、点線ルートでなく実線の最少消費電力ルートが選択されている。すなわち、ネットワークとサーバを含めた全体での消費電力が最小となるルート選択の実現を確認した。

4.4.2 SLAを保証した省電力サーバ動的配置

メトロネットワーク/アクセスネットワークで接続された図7に示す神奈川DCと宮城DCを用いてサンプルWebサービスを運用し、提案手法の検証を行った。マッピングノードの構成には、オーバーレイネットワーク構築ツールキットOverlayWeaver[16]を用いた。両DCで稼働するマッピングノード群が、SLA違反をしない範囲で省電力効果の高いDCにリクエストを転送することを確認した。

5. おわりに

本論文では、大規模クラウドシステムにおいて、提供するサービスの品質を劣化することなくシステム全体の消費電力を削減する手法を提案した。宮城、東京、神奈川

の3都県に模擬的に配備した複数のDCをアクセスネットワークで接続し、さらにアクセスネットワークを上位のメトロネットワークで接続した検証用の大規模な省電力クラウドシステムを試作した。初期条件として、DCに合計1,000VMを配備した設定において、サービス品質を維持したVMの片寄せがメトロネットワークを跨いで可能なこと、およびシステム全体として20%から30%の省電力化が可能なことを検証した。また、メトロネットワークでは、VM移動に伴うメトロネットワークの伝送経路修正を1秒間に約250のVM移動要求に対応可能なことを検証した。さらに、複数サーバ上に配備したサービスパーツを連結提供するマッシュアップサービスの評価では、消費電力が最小となる経路上にあるサービスパーツを連結したマッシュアップ機能を検証した。そして、ユーザからのリクエスト処理では、SLA違反をしない範囲で、省電力効果の高いDCにおけるサーバ動的再配置処理を検証した。システム全体として、ネットワークとサーバとが連携することにより、サービスの品質を劣化することなく、サービス提供が可能な省電力クラウドシステムの実現が可能なことを検証した。

今後は、より実サービスシステムに近い構成や規模へ対応するためのエンハンスを図る予定である。

謝辞 本研究の一部は、総務省の委託研究「広域災害対応型クラウド基盤構築に向けた研究開発(環境対応型ネットワーク構成シグナリング技術)」の一環として実施された。

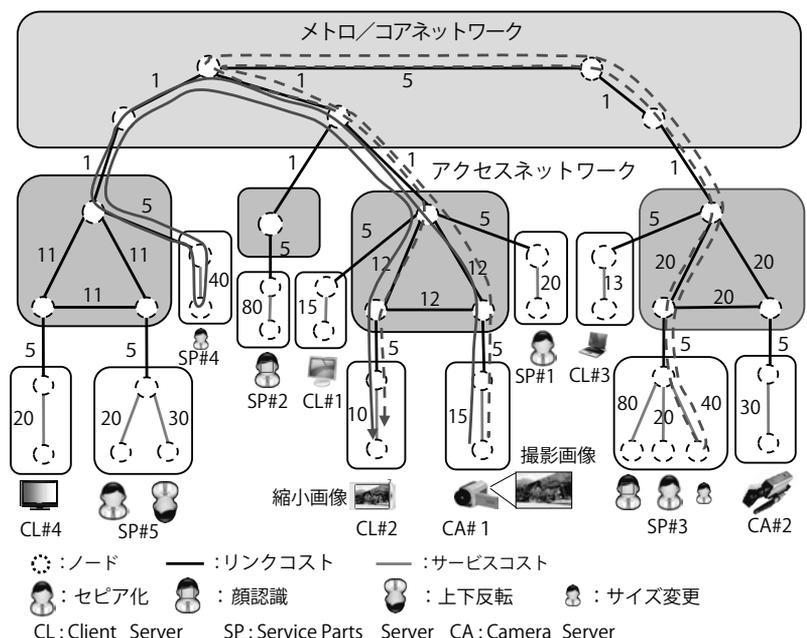


図11 実証実験システムのサービスレイヤトポロジ

参考文献

- 1) Belady, C.L.: Microsoft Corporation, Projecting Annual New Datacenter Construction Market Size (Mar. 2011).
http://cdn.globalfoundationservices.com/documents/Projecting_Annual_New_Data_Center_Construction_PDF.pdf [retrieved: May, 2013]
- 2) GreenTouch, <http://www.greentouch.org/> [retrieved: May, 2013]
- 3) Yang, S., Chen, L., Tseng, H., Chung, H. and Lin, H.: Designing Automatic Power Saving on Virtualization Environment, IEEE International Conference on Communication Technology (ICCT 2010), pp.966-970 (Nov. 2010).
- 4) Yamada, M., Yazaki, T., Matsuyama, N. and Hayashi, T.: Power Efficient Approach and Performance Control for Routers, Proc. of IEEE International Conference on Communications (ICC Workshops 2009), pp.1-5 (June 2009).
- 5) Zhang, Y., Chowdhury, P., Tornatore, M. and Mukherjee, B.: Energy Efficiency in Telecom Optical Networks, IEEE Communications Surveys & Tutorials, Vol.12, No.4, pp.441-458, Fourth quarter (2010).
- 6) Shimizu, S., Imai, S. and Yamada, A.: Auxiliary Graph Based Approach for Impairment Aware Green Lightpath Provisioning, 17th Asia-Pacific Conference on Communications (APCC 2011), pp. 615-620 (Oct. 2011)
- 7) <http://www.openflow.org/> [retrieved: May, 2013]
- 8) <http://www.projectfloodlight.org/floodlight/> [retrieved: May, 2013]
- 9) <http://openswitch.org/> [retrieved: May, 2013]
- 10) Nakahara, K., Kikuta, K., Yamada, S., Ishii, D., Okamoto, S. and Yamanaka, N.: Realizing the Scalable Service Provisioning in Ubiquitous Networking Environment (uGrid) by Extending Routing Protocol, IEICE Technical Report, Vol.NS2010, No.56, pp.49-54 (Sep. 2010).
- 11) Shibuta, N., Nakahara, K., Kikuta, K., Ishii, D., Okamoto, S., Oki, E. and Yamanaka, N.: Service Composition System Optimizing Network and Service Resources in E3-DCN, IEICE Technical Report, Vol.111, No.344, NS2011-138, pp.115-120 (Dec. 2011).
- 12) Shibuta, N., Zhang, S., Takeshita, H., Okamoto, S., Oki, E. and Yamanaka, N.: Dynamic Transmission Route Selection Towards the Contents Transmission Energy Optimization in E3-DCN, IEICE Technical Report, Vol.112, No.208, NS2012-70, pp.99-104 (Sep. 2012).
- 13) Qureshi, A., Weber, R., Balakrishnan, H., Gutttag, J. and Maggs, B.: Cutting the Electric Bill for Internet-Scale Systems, Proc. of the ACM SIGCOMM 2009 Conference on Data Communication, pp.123-134 (Aug. 2009).
- 14) Heller, B., Seetharaman, S., Mahadevan, P., Yiakoumis, Y., Sharma, P., Banerjee, S. and McKeown, N.: ElasticTree: Saving Energy in Data Center Networks, Proc. of the 7th USENIX Conference on Networked Systems Design and Implementation, pp.17-17 (Apr. 2010).
- 15) Stoica, I., Morris, R., Karger, D., Kaashoek, M.F. and Balakrishnan, H.: Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications, Proc. of ACM Special Interest Group on Data Communications Conference, pp.149-160 (Aug. 2001).
- 16) Shudo, K., Tanaka, Y. and Sekiguchi, S.: OverlayWeaver: An Overlay Construction Toolkit, Computer Communications, Vol.31, No.2, pp.402-412 (Feb. 2008).
- 17) Common Open Research Emulator Project, <http://cs.itd.nrl.navy.mil/work/core/index.php> [retrieved: May, 2013]

鈴木 敏明 (非会員) toshiaki.suzuki.cs@hitachi.com

1990年東京理科大・理工学部卒。1992年同大学院修士課程了。同年(株)日立製作所入社。以来、臨場感映像通信システム、

アクティブネットワーク、省電力クラウドシステムの研究開発に従事。

垂井 俊明 (正会員) toshiaki.tarui.my@hitachi.com

1985年東大・工学卒。1987年同大学院修士課程了。同年(株)日立製作所入社。以来、並列計算機、自律管理システム、仮想化、省電力クラウドの研究開発に従事。

馬場 智宏 (非会員) tomohiro.baba.mn@hitachi.com

1997年東京工科大・工学部卒。1999年同大学院修士課程了。同年(株)日立製作所入社。以来、ルータの設計開発、ネットワーク関連のSEに従事。

高瀬 晶彦 (非会員) akihiko.takase.wa@hitachi.com

1981年東大・大学院理学系研究科修了、理博。同年(株)日立製作所入社。以来、光アクセスシステム、ATMシステム、フレームリレー、IPサービスシステム、ネットワークセキュリティ、パケット・光トランスポートシステムの研究開発および事業開発に従事。

山田 亜紀子 (非会員) akikoo@jp.fujitsu.com

1993年お茶の水女子大・理学部卒。1995年同大学院修士課程了。同年日本経済新聞社入社、新聞製作系システムの企画導入に従事。2001年(株)富士通研究所入社。以来、ネットワーク設計制御技術の研究開発に従事。

清水 翔 (非会員) shimizu.sho@jp.fujitsu.com

2007年慶大・大学院修士課程了。2010年同大学院博士課程了、博士(工学)。同年(株)富士通研究所入社。以来、省電力ネットワーク、SDNの研究開発に従事。

宗宮 利夫 (非会員) soumiya.toshio@jp.fujitsu.com

1987年佐賀大・理工学部卒。1989年同大学院修士課程了。同年(株)富士通研究所入社。局用交換システムのトラヒック制御機能開発、マルチレイヤネットワーク制御、省電力ネットワーク設計制御の研究開発に従事。

竹下 秀俊 (非会員) takeshita@yamanaka.ics.keio.ac.jp

1974年静岡大・工学部卒。同年日本電気(株)入社。電子交換機、デジタル交換機の研究開発、固定電話通信事業者、および携帯電話通信事業者向けSE業務に従事。2012年慶應大大学院博士課程了、博士(工学)。省電力ネットワーク、省電力サービスの研究開発に従事。

堀江 光 (学生会員) hikaru.horie@ssl.ics.keio.ac.jp

2009年慶大・理工学部卒、2011年同大学院修士課程了。同年同博士課程に進学、分散システムおよびデータセンタの省電力化手法の研究開発に従事。

山中 直明 (正会員) yamanaka@ics.keio.ac.jp

1983年慶大・大学院修士課程了、同年日本電信電話公社入社、1991年博士(工学)取得、2004年慶應義塾大学教授。超高速ATMスイッチ、IPバックボーンルータ、フォトニックネットワーク、GMPLSネットワーク、光アクセスシステム、スマートグリッド、新世代ネットワークの研究開発に従事。電子情報通信学会フェロー、IEEEフェロー、前IEEE Asia Pacificディレクタ、2003～2013年度けいはんな情報通信オープンラボ相互接続性検証ワーキンググループ主査、電子情報通信学会編集理事。

投稿受付: 2013年5月22日

採録決定: 2013年7月26日

編集担当: 丸山 宏 (統計数理研究所)