

平成 30 年 5 月 24 日現在

機関番号：11301

研究種目：挑戦的萌芽研究

研究期間：2014～2017

課題番号：26540032

研究課題名(和文)ビッグデータ指向型ネットワーク：複雑ネットワーク理論に基づく物理と仮想の融合

研究課題名(英文)The big data oriented network: Fusion of physical and virtual based on complex network theory

研究代表者

加藤 寧 (KATO, NEI)

東北大学・情報科学研究科・教授

研究者番号：00236168

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)： ネットワーク上で生成される膨大なデータを利用・解析し、革新的なサービスやビジネスモデルを創出するビッグデータ解析が注目されている。ビッグデータ解析では、アクセス網やコア網、データセンタ網などネットワーク全体の資源の効率的な利用が必要とされており、ネットワーク全体を最適化する広域ネットワーク仮想化技術の実用化が望まれている。しかしながら、広域ネットワーク仮想化に関する議論は始まったばかりであり、未だ確固たる技術は開発されていない。そこで本研究は、複雑ネットワーク理論に基づく広域ネットワーク仮想化技術を開発し、ビッグデータ解析の要求品質を保証可能なビッグデータ指向型ネットワークを創出した。

研究成果の概要(英文)： Big data analysis that uses and analyzes the vast amounts of data generated on the network and creates innovative services and business models has attracted recent attention. Big data analysis requires efficient use of resources on the entire network to be able to analyze access networks, core networks, and data center networks. And practical application of wide area network virtualization technology to optimize the entire network is desired. However, the discussion on wide area network virtualization has just begun, and yet no effective technology has been developed. Therefore, in this research, we developed wide area network virtualization technology based on complex network theory to create big data oriented network which can guarantee required quality of big data analysis.

研究分野：総合領域

キーワード：ビッグデータ ネットワーク仮想化 複雑ネットワーク理論 分散処理 広域ネットワーク

## 1. 研究開始当初の背景

インターネットが社会的に重要なインフラとなり、伝送制御プロトコル (Transportation Control Protocol : TCP)/インターネットプロトコル (Internet Protocol : IP) が世界標準のネットワークアーキテクチャとして開発された。近年では、その高い有用性から、TCP/IP のネットワークが情報通信技術 (Information and Communication Technology: ICT) サービスの提供にも使用されている。しかしながら、TCP/IP のネットワークはエンドツーエンドの通信の安定性に特化しているために、サービスの提供に必要な要件に対応しきれていない。このサービスに必要な要件とネットワークで可能なこととの間のギャップがあるためにユーザのサービスへの満足度が低下している。したがって、ICT サービスを基盤としたスマート社会の実現のため、TCP/IP ネットワークの改善が大きな課題となっている。

Content-Centric Networking (CCN) や Information-Centric Net-working といった新世代のネットワークアーキテクチャには仮想ネットワーク技術が使用されている。これらの新たなアーキテクチャでは IP ネットワークに代わって仮想ネットワークが使用され、任意の端末間のデータ通信を可能にしている。加えて、物理ネットワークのリソースを適切に使用することで、様々な要求を満たす最適な仮想ネットワークを構築することができる。さらに、仮想ネットワークによって、物理ネットワーク内に複合的な仮想ネットワークを構築することが可能であり、この構築したネットワークを使用することで、物理ネットワーク内で様々な ICT サービスを提供することが可能となる。このような変化は ICT サービスの統合や今後のビッグデータを使用する社会の構築に重要である。

新世代のネットワークアーキテクチャでは、ネットワーク全体のリソースを使用した仮想ネットワークを構築する。しかしながら、ネットワーク全体の仮想化によって、発生するネットワークの障害が複雑化してしまう。よって、通信端末の故障や通信経路の障害、災害などを考慮に入れなければならない。上述したネットワーク障害は仮想ネットワークの性能に多大な影響を及ぼす。本研究では、フォールトトレラントな仮想ネットワークを調査する。このネットワークでは想定しうるあらゆるネットワーク障害が発生しても、高い通信性能を維持可能である。しかし、世界的にもフォールトトレラント性を持った仮想ネットワークの研究開発は最近に開始されたものである。そのため、多くの研究課題が存在する。

## 2. 研究の目的

社会的な背景に応えるにあたって、仮想ネットワークにはフォールトトレラント性に関する課題が存在する。本研究の狙いはフォ

ールトトレラントな仮想ネットワークを構築する方式の開発である。本研究で我々は以下のネットワーク障害に焦点を当てた。

### (1) 仮想ネットワークの障害 (Failures in Virtual Network: FVN)

FVN は、ソフトウェアのバグや DoS 攻撃によって仮想マシン (Virtual Machines: VMs) が機能を停止することである。こういった障害のために、故障した VMs は仮想ネットワークから取り除かれ、仮想ネットワークが分断される。その結果、まだ機能している VMs が孤立する。したがって、FVN が発生しやすい環境では仮想ネットワークのコンピュータ処理と通信の性能の両方が低下する。上記の課題を解決するために、本研究では仮想ネットワークの最適化に焦点を当てる。

### (2) 物理ネットワークの障害 (Failures in Physical Network: FPN)

FPN は通信処理機器 (例: ルータ、スイッチ、サーバ、ユーザの使用する端末) の故障と物理ネットワークの障害を含む。以上の障害が原因となって、VM がルータによって収容される、あるいはサーバ上で起動された VM が機能を停止することで、使用可能な VM 数が減少する。この課題を解決するために、物理ネットワークのトポロジ情報を考慮して、仮想ネットワークを構築する手法を改善することが本研究の目的である。

我々の研究目的には前述した ICT サービスを仮想ネットワークでの利用を拡大することも含んでいる。現在、ビッグデータマイニングサービスが大きな注目を集めており、今後の ICT 社会でますます重要となる。しかし、近年普及しているアーキテクチャは依然として課題が存在する。この課題に取り組むため、本研究では完全に分散したビッグデータマイニングを提案する。

以上のように本研究は現在存在するネットワークアーキテクチャに代わる、耐障害の仮想ネットワークを提案する。これは、次世代のネットワークアーキテクチャに対して耐障害の点で大きく寄与するものである。

## 3. 研究の方法

音声通話や Web といった通常のサービスでは、端末と端末やサーバと端末の通信が主なトラフィックである。この際、ユーザは特定の端末を探索する必要がある。つまり、通常のサービスを利用するために、ユーザは端末の位置を認識する必要がある。これは端末指向のサービスと言える。したがって、一般的には常に IP に基づくネットワークを利用しており、それぞれの端末は接続したい端末を IP アドレスによって検索する。これがインターネットの基本である。その一方でユーザはインターネットを介して ICT サービスに向けて特定のコンテンツを提供し、同様に

獲得する．これはコンテンツの送信とビッグデータマイニングのようなコンテンツ指向のサービスである．したがって，通常の IP に基づくネットワークはコンテンツ指向のサービスと新たなアーキテクチャに適していない．これらの高い性能を持つサービスには，コンテンツやデータの情報に基づいて検索を行うメカニズムが必要である．

将来的にはインターネット上で通信するコンテンツが爆発的に増加することが予想される．したがって，非常に大量のコンテンツやデータを高い効率かつ低いコストで提供，取得できる技術開発が必要である．以上の理由から，本研究では仮想ネットワークを使用する検索アーキテクチャを提案する．

図1に提案する仮想ネットワークを使用した検索アーキテクチャを示す．このアーキテクチャでは，仮想ネットワークを使用することで，ユーザ，あるいは VM が求めるデータを検索する．データを見つけるため，それぞれの VM は隣接する VM の IP アドレスを把握している．データを発見した時，それぞれの VM(クライアント VM)は検索クエリを隣接する VMs に送信する．このクエリには求めるデータの ID とその IP アドレスが記載されている．クエリを受信した VMs が要求されたデータを所持していなかった場合，自身の隣接 VMs にクエリを送信する．この動作は求めるデータを発見するまで続く．求めるデータを所持する VM が検索クエリを受信した時，その VM はクライアント VM に向けて物理ネットワークを介してデータを送信する．

従来の IP に基づくアーキテクチャでは，中央の VM がデータの位置を保管していた．つまり，それぞれのクライアント VM がデータを持つ VM の IP アドレスを中央 VM に問い合わせているのである．しかし，ネットワーク内に大量のデータが存在するとき，中央の VM はデータの位置を定期的に取得する必要があるため，このアーキテクチャではサービスの可用性が低下する．これとは対照的に，本研究で提案するアーキテクチャでは，それぞれの VM が隣接する VMs の IP アドレスのみを保管し，データの位置を知る必要がない．したがって，データやコンテンツ量が増加しても提案するアーキテクチャは高い可用性を維持できる．

#### 4. 研究成果

本研究では仮想ネットワークを使用した検索アーキテクチャの性能を高める技術として，仮想ネットワークトポロジ，VM マッピング戦略，動的変更システムを用いる．

##### (1)仮想ネットワークトポロジ

仮想ネットワークトポロジは提案する検索アーキテクチャの性能に影響を及ぼす．物理ネットワークの情報に関わらず仮想ネットワークのトポロジは決定される．これは，

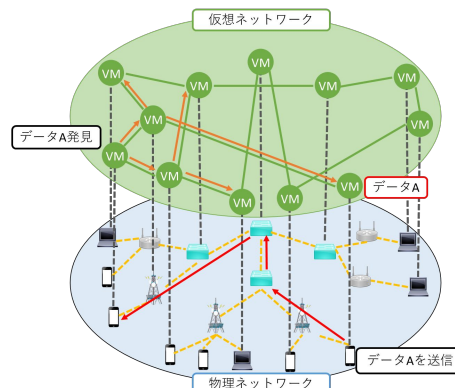


図1.想定ネットワーク

サービスのプロバイダがそのサービスの要求や目的を満たす仮想ネットワークトポロジを構築することを意味する．現在の仮想ネットワークトポロジを使用した検索技術では，検索の効率を高めるために平均ホップ数を削減している（あるいは，仮想ネットワークの直径を縮小する）．検索効率を高める基本的な考え方は，それぞれの VM のリンク数を増加させることである．しかしながら，このアプローチには短所が2点ある．1点目は仮想ネットワークの構築に長時間必要になること．2点目は全ての VM に検索メッセージを送信することでトラフィックが爆発的に増加してしまうことである．したがって，実際に規模の拡大が可能なサービスの提供するためには，より小規模な処理を行わなければならない．

##### (2)VM マッピング戦略

VM マッピングとは仮想ネットワーク内の VM として物理ネットワークの VM を選択することである．つまり，VM マッピングが物理ネットワークと仮想ネットワーク間の関係を決定しているため，検索アーキテクチャの性能に大きな影響を及ぼす．仮想ネットワークトポロジを考慮することでサービスのプロバイダはサービスの要求を満たそうとしているにも関わらず，検索性能（検索速度や成功率）物理ネットワークの通信リンクとデバイスの性能（接続の速さ，通信容量，処理速度）の影響を受けてしまう．したがって，要求を満たすために，VM マッピング戦略をデザインする必要がある．

VM マッピング戦略には二つの種類がある．一つ目は(i)与えられたリソースに基づく VM マッピング戦略である．二つ目は(ii)動的なリソースに基づく VM マッピングである．(i)の場合は VM の通信容量と VM 同士の接続があらかじめネットワーク事業者によって設定されている．そして，サービスのプロバイダ与えられた情報に基づいて VM マッピングを最適化する．もう一方で(ii)ではサービスのプロバイダとネットワーク事業者が協力してサービスを提供するか，ネットワーク事業者がサービスを提供する．この状況ではサービスのプロバイダは物理リソースと VM の性能や接続の利用量を変更する要求を送信す

る．したがって、サービスのプロバイダはサービスの性能を高めるために VM マッピング戦略とリソース割り当てを統合する必要がある．

### (3)動的な変更システム

サービスの性能の要件はユーザの要求を満たすよう状況によって刻一刻と変化する．物理的な VM の性能と接続は様々なネットワーク環境から変化する(例: VM の故障)．したがって、ネットワークの状況や環境が変化したとき、サービスの要求を満たすために仮想ネットワークポロジと VM マッピングの両方を使い分ける必要がある．

加えて、「認知」と「知能」の機能が上記の二つの技術の動的な変更の実現に必要である．仮想と物理、両方のネットワーク環境と状況を観測して、情報を収集するのが「認知」機能である．その一方で、収集した情報をマイニングすることで仮想ネットワークの状態と VM マッピングをリアルタイムでの最適化するのが「知能」の機能である．現在の状態が最適な状態と異なるとき、サービスプロバイダは仮想ネットワークの変更を試みる．

上記の技術からなる仮想ネットワークを利用した検索アーキテクチャは高い性能を発揮して、要求を満たすにも関わらず、物理ネットワークでの FVN と FPN によって VM は機能を停止してしまう．機能を停止した VM が仮想ネットワークから取り除かれ、いくつかのまだ機能している VM が仮想ネットワーク内で孤立してしまう．つまり、まだ使用できる VM が隣接する VM との仮想的な接続を失ってしまう．この状況では、検索アーキテクチャは孤立した VM を発見できず、サービスの可用性は著しく低下してしまう．また、VM が機能を停止することで、VM 間のホップ数が増加して、検索効率が低下することもある．VMs が機能を停止した時が挙げられる．以上より、ネットワーク耐障害の課題を対処する必要があり、この課題の解決は高性能な仮想ネットワークを使用した検索アーキテクチャの実現に必要不可欠である．

従来の VM がデータマイニング機能を管理するアーキテクチャでは障害への耐性がないが、本研究で提案する新たなビッグデータマイニングアーキテクチャでは、データマイニング機能は分散しており、仮想ネットワークを利用することで管理されている．加えて、本研究では FVN と FPN に耐性を持つ仮想ネットワークを構築する手法を提案する．また、FPN が発生した後のサービスの可用性の観点で今回提案するビッグデータマイニングアーキテクチャの有用性を示す．

Hadoop は最も一般的なビッグデータマイニングサービスを提供するアーキテクチャ

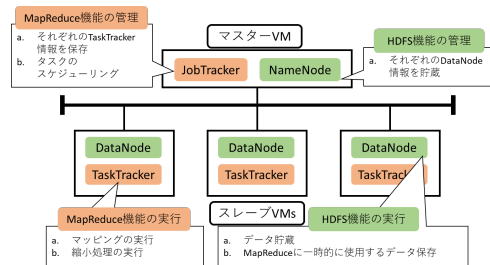


図 2. Hadoop アーキテクチャ

である．このアーキテクチャでは処理を行う VM の数に比例してデータマイニングの速度が上昇する．図 2 は Hadoop アーキテクチャを示している．この図からわかるように、Hadoop は MapReduce と Hadoop Distributed File System (HDFS) からなっている．

MapReduce は、複数のノードを使用して並列分散アルゴリズムで大量のデータセットを処理および生成する．HDFS は MapReduce で使用されるビッグデータを管理する．さらに、処理を行う VM は二つに分類される．一つのマスターVM と複数のスレーブ VM である．マスターVM はビッグデータマイニングを管理するために JobTracker 機能と NameNode 機能を持つ．JobTracker はそれぞれの TaskTracker の情報を管理し、その情報に基づいてタスクのスケジューリングをする．NameNode はデータベースを持ち、求めるデータを検索する．反対にそれぞれのスレーブ VM は二つの機能を持つ．一つは MapReduce を実行する TaskTracker、二つ目は HDFS にデータを格納する DataNode である．

Hadoop では処理要求が入力されると、マスターVM は HDFS を使用して入力された処理要求に属するデータを見つける．つまり、NameNode がデータベースを検索することで、要求されたデータをもつスレーブ VM を見つけようとする．マスターVM がスレーブ VM からデータを受信した後、マスターVM の JobTracker がマッパーを決定する．このマッパーとは、それぞれの TaskTracker の処理負荷に基づいてマッピング処理を行うスレーブ VM である．マッパーの選択に続いて、JobTracker はデータをいくつかに分割して、選択したマッパーに割り当てる．ここで、データの分割数はマッパーの数に基づいて決定する．マッパーが JobTracker から分割されたデータを受け取ると、マッパーは大量の情報をマッピング処理で分類分けする．そして次の縮小フェーズで必要となる情報を抜き出す．全てのマッパーがマッピング処理を終えると、JobTracker はデータ縮小役の VM を決定する．縮小役はデータ縮小処理をおこなうスレーブ VM である．縮小役はマッパーから選ばれる．縮小役はマッパーからマッピング処理を行った情報を抽出して収集する．そして、縮小役は収集した情報を結果にまとめる縮小処理を行う．最後に、縮小役は獲得した結果を NameNode に送信し、そしてスレーブ VM はデータを格納する．

Hadoop はデータマイニング処理を行う端末数に比例した速度で行うことが可能となる．その性能はタスクの割り当てとスケジューリング方式に依存している．したがって、処理速度、ネットワークリソースの効率性、処理リソースの効率性、消費エネルギー効率を高める高性能な並行データマイニングアーキテクチャが開発されている．Hadoop には大きなメリットがあるにも関わらず、VM によるネットワーク障害の影響を受けやすい．これはハードウェアの問題やソフトウェアのバグのために VM が機能しなくなると、データマイニングの成功確率とマイニング速度は大幅に低下するためである．この耐障害的な課題を解決するため、一般的な Hadoop アーキテクチャでは複製を作るというアプローチを使う．この手法によって、冗長タスクを他の VM に割り当てることで、VM の機能停止に対する可用性を高めることが可能である．しかし、この手法をもってしても現実的な環境で十分なサービスの可用性を保証することは困難である．理由は、故障の規模や発生確率によって、最適な複製の数やマスター VM の数が変化するためである．このようにネットワーク障害における課題に取り組んでいるが、想定している VM の故障は小規模なものである．つまり、既存の提案されている技術では大規模な障害は考慮していないため、FPN といった大規模な障害が発生するとデータマイニングの性能が著しく低下する．したがって、将来のユビキタスビッグデータマイニングサービスのために、FPN への耐性を持つデータマイニング技術が必要不可欠と言える．

#### (4) 提案するビッグデータマイニングアーキテクチャ

従来の中央管理を使用するビッグデータマイニングアーキテクチャは以下の短所がある．(i) 単一点障害のようなマスター VM が機能を停止する状況では管理不能となる点．(ii) VMs の数が増加すると性能が大幅に低下し、サービスが停止する点．これらの課題を解決するため、本研究では仮想ネットワークを使用する並行ビッグデータマイニングアーキテクチャを提案する．このアーキテクチャでは JobTracker や NameNode といった管理機能が全ての VM で分散して行われる．マスター VM に代わって、仮想ネットワークが全ての VM で構築され、データと処理中の VM の検索に使用される．全ての VM が管理機能と処理機能の両方を実行するため、複数の VM が故障した場合や VM 数が増加した時でもこのアーキテクチャは高い性能を発揮する．つまり、提案するアーキテクチャは仮想ネットワークが断絶された場合でも性能の高いビッグデータマイニングサービスを提供可能である．

図 3 は本研究で提案するネットワーク内全ての VM からなる仮想ネットワークを使用し

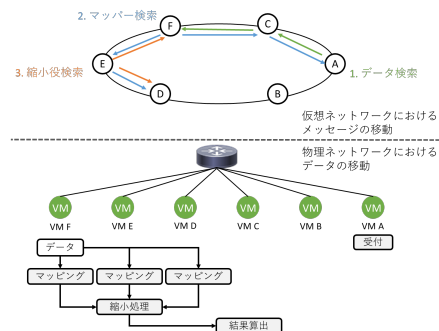


図 3. 仮想ネットワークを使用したアーキテクチャの例

たアーキテクチャの例とそのアーキテクチャが MapReduce と HDFS を実行する様子を示している．データマイニングの要求が入力されると、クライアントから要求を受け取った受付 VM が受け入れ処理を行う．この受け入れ処理では、仮想ネットワークを使用して、受付 VM が入力されたマイニング要求に関連するデータを持つ VM (あるいは VMs) を検索する．この検索される VM (あるいは VMs) は所有者と呼ぶ．ここで使用する検索方式としては、フラッディング、ハッシュテーブルなど複数存在する．次に所有者(例: 図 3 の VM F) は自身の CPU 使用率に基づいてデータの分割数を決定し、データマイニングに使用するデータを分割する．さらに、所有者は仮想ネットワークで近い端末からマッパーを選択する．ここで VMs の中で CPU 利用率が特定の閾値以上のものは選択されない．例を挙げると、図 3 では VM D, E, F がマッパーとして選択される．マッパーは所有者から送信された分割されたデータを使用してマッピング処理を実行する．そして、最も早くマッピング処理を完了したマッパーは縮小役(例: 図 3 VM E)となる．処理の完了したデータを縮小役に送信するよう要求するメッセージを縮小役は他のマッパーに送信する．そして、全てのマッパーから処理後のデータを受信した後に縮小処理を行う．最後に縮小者は得た結果を格納する．

上記のように仮想ネットワークを使用し、全ての管理機能と処理機能が分散して実行される．本研究で提案するアーキテクチャを使用することで、従来のアーキテクチャの単一点障害は解決することができる．しかし、提案するアーキテクチャのサービス可用性は仮想ネットワークの接続性の影響を受けてしまう．したがって、ビッグデータマイニングのための耐障害仮想ネットワークが必要である．

#### (5) 仮想ネットワークの構築方式がサービス可用性に及ぼす影響

ここでは、提案した物理ネットワーク情報に基づく仮想ネットワークを構築する方式とランダムにネットワークを構築する方式のそれぞれについて、データマイニングアーキテクチャのサービス可用性を評価する．

物理ネットワークポロジはツリー構造であり、仮想ネットワークの平均規模は 3, VM の数は 10,000 と仮定する。さらに実行されるビッグデータマイニングのタスク数は 1000 個, それぞれのタスクに関連するデータ数と複製の数はそれぞれ 5 個と 2 個と仮定する。以上の設定で, FPN が発生した後のそれぞれのタスクの成功率を算出することで, サービスの可用性を評価する。ここで規模の異なる FPN を想定する。順に小規模, 中規模, 大規模の FPN があり, これらの FPN が発生するとそれぞれ約 100, 1000, 3000 の VM が他の VM と接続不可となる。

図 4 は 3 つの規模の FPN が発生した場合の提案した方式のデータマイニングサービスにおける可用性を評価している。結果からわかるように, 提案した方式では FPN の規模に関わらず, 比較手法より高いサービス可用性を発揮しており, その値はおよそ 100% となっている。この結果の理由としては, 提案方式では, 仮想ネットワークが断絶されず, 全てのタスクに関わる全てのデータが大規模クラスタ内に存在するためである。その一方で, ランダムにネットワークを構築する方式では, 機能しなくなった VM 数が増加するに従って, サービスの可用性が低下した。この理由は機能を停止した VM が増加することで孤立する VM が増加するためである。

#### (6) データ配置方式がサービス可用性に及ぼす影響

FPN が発生後のサービス可用性の観点で提案するデータ配置方式の効果をランダムにデータを配置する方式と比較することで検証する。ここでは仮想ネットワークが提案した方式に基づいて構築されている。

図 5 はデータマイニングのサービス可用性について, 提案したデータ配置方式の効果を示している。評価では前述したものと同様の設定を想定している。既存のデータ配置方式ではたとえ高い冗長性を持っていたとしても, 機能を停止する端末が徐々に増加していくにつれて, 急激にサービス可用性が低下する。その一方で, 提案した配置方式では, 大規模クラスタ内でタスクに付随するデータを保持する VM の存在を保証するため, 機能を停止する端末の数に関わらず, 最小の冗長性でデータマイニングの成功確率 100% を達成している。以上から, 提案したネットワーク構築方式とデータ配置方式を備えたビッグデータマイニングアーキテクチャは, FPN が発生した場合でも高い成功率のデータマイニングサービスを提供可能だと結論付けられる。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

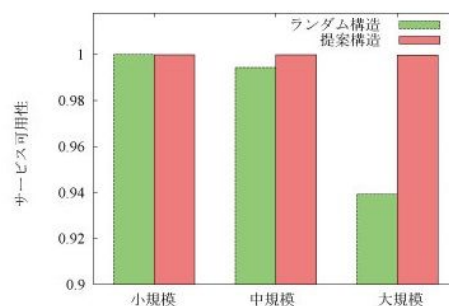


図 4. 仮想ネットワークの構築方式がサービス可用性に及ぼす影響

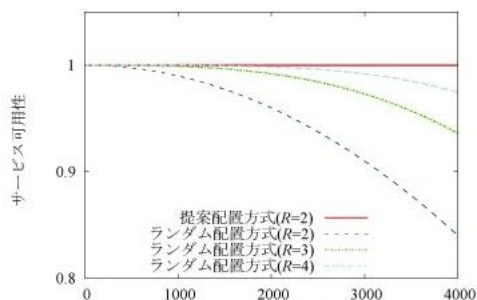


図 5. データ配置方式がサービス可用性に及ぼす影響

- (1) Shikhar Verma, Yuichi Kawamoto, Zubair Md. Fadlullah, Hiroki Nishiyama and Nei Kato, "A Survey on Network Methodologies for Real-Time Analytics of Massive IoT Data and Open Research Issues," IEEE Communications Surveys and Tutorials, 査読有, vol. 19, no. 3, pp. 1457-1477, Apr. 2017, DOI: 10.1109/COMST.2017.2694469.

[学会発表](計 1 件)

- (1) 高木啓彰, 西山大樹, 加藤寧, "リアルタイム性を考慮した IoT システムの実現に向けた フィードバック制御に関する一考察," 電子情報通信学会 2016 年総合大会, 2016.

[その他]

ホームページ等

<http://www.it.is.tohoku.ac.jp/>

#### 6. 研究組織

##### (1) 研究代表者

加藤 寧 (NEI KATO)

東北大学・大学院情報科学研究科・教授

研究者番号: 00236168

##### (2) 研究分担者

なし

##### (3) 連携研究者

なし

##### (4) 研究協力者

なし