

新世代ネットワーク構想における ネットワーク仮想化

Research on Network Virtualization for New Generation Networks

中尾彰宏



情報爆発時代を迎え、あらゆる情報に対するネットワークの利用方法が多岐多様になる中、ネットワークを自由な発想で白紙から創造・設計することを可能とし、更にそれらを複数同時に収容することを目的とする情報社会基盤の構築が望まれている。こうした中、ネットワーク仮想化技術の進化により、高速通信を維持したまま多様で柔軟なプログラマビリティやアクセス制御機能を実現するネットワーク基盤が実現されつつある。本稿では、全く新しいネットワークの構築・検証や、クラウドサービス等の新しいサービスの実現を可能とするための「進化した」ネットワーク仮想化環境の実現を目指した取組みについて論じる。将来のネットワークはどのようにあるべきかという観点にも触れる。

キーワード：ネットワークアーキテクチャ、ネットワーク仮想化、クラウドネットワーク

1. はじめに

世界各国で、近年、本来あるべき姿のネットワークを「白紙」から再設計する研究が進められている。我が国でも総務省が主導して新世代ネットワークの研究が早い時期を迎えている。

我々は、これまで、既存のネットワーク技術に追加的（インクリメンタル）に新しい機能を付加したり改良を加えたりする研究開発を遂行してきた。しかし、セキュリティの問題、ネットワーク接続の可用性や堅ろう性の問題など、多数の問題が従来の研究開発方法では解決が困難であることが指摘されている。このため、新世代のネットワークは白紙から設計（クリーンスレートデザイン）し、様々な新しい通信手順を提案・実装・検証する研究が、世界各国で始まっている。

しかし、従来のネットワーク基盤では、新しいネットワーク機能が創造されても、実際のネットワークに実装するまでには数～10 数年かかる場合が多く、ネットワーク基盤の劇的な進化が困難であるという現実的な限

界が指摘されてきた⁽¹⁾。また、同時に、新しく提案されつつある多種多様なネットワークアーキテクチャが混在する中、最良な単一のネットワークアーキテクチャに収束させ、それを新世代のネットワークとして採用するか、あるいは、複数のネットワークが混在する状態を本来あるべき通信基盤の姿とするかは議論を要する問題である。これらの問題を解決するには、複数のネットワークを同時に収容し、実証実験を可能とすることが必要である。つまり、共通のネットワークインフラ全体を論理的に分割し、論理ネットワーク上で自由なネットワーク機能の創造を可能にする環境が求められている^{(2), (3)}。

2. 将来のネットワークビジョン

我々は、近い将来、多様なネットワークが普遍的に存在すると予見し、また、こうしたネットワークを複数同時収容するためには現在のネットワーク基盤では限界があると考えている。この予見は、以下のように、学術研究界やビジネス社会の動向に裏付けされている。

第1に、クラウドコンピューティングを推し進める米国巨大企業は、我々学術界の研究者よりもいち早く新しいネットワークの仕組みやプロトコルを考案し実装する可能性がある。

第2に、ネットワークに対する要件の異なるアプリ

中尾彰宏 正員 東京大学大学院情報学環
E-mail nakao@iii.u-tokyo.ac.jp
Akihiro NAKAO, Member (Interfaculty Initiative in Information Studies, The University of Tokyo, Tokyo, 113-0033 Japan).
電子情報通信学会誌 Vol.94 No.5 pp.385-390 2011年5月
©電子情報通信学会 2011

ケーションが増加しつつある。ピアツーピアトラフィックや、クラウド型の動画配信トラフィックが増加している一方で、医療ネットワークや防衛セキュリティのためのネットワークなど、セキュリティレベルの非常に高いネットワークも求められている。

第3に、前述のように学術研究界では、従来のネットワークに捉われない新しいネットワークの研究が始まっている。米国では、GENI/FIND^{(3)~(5)}などのプロジェクトを推進し、欧州ではFP7⁽⁶⁾というプロジェクトを推進している。諸外国では、大学院生が新しいネットワークの仕組みを提案することが日常化している。

3. 進化するネットワーク仮想化技術

3.1 進化するネットワーク仮想化技術の定義

我々は、「広義のネットワーク」を、データを伝搬する「リンク」とプロトコルを解釈しリンクを選択するプログラムを実行する「ノード」を基本要素として構成される「情報通信基盤」として定義する。例えば、現在のインターネットにおいては、イーサネットや光ファイバなどはデータパケットを伝搬する「リンク」であり、ルータやスイッチは、データパケットのヘッダ情報をプロトコルに基づき解釈し、次に移動するべきルータやスイッチあるいはエンド端末へのリンクを選択するプログラムが実行される「ノード」である。

新世代における「ネットワーク仮想化技術」とは、この広義のネットワーク全体を仮想化する、つまり、リンク資源だけではなくノード上の計算資源やストレージ資源も含むインフラストラクチャ全体の資源の仮想化と定義される。ルータやスイッチは、CPUやメモリを計算資源とストレージ資源として保持する一種の計算機にすぎない。クラウドコンピューティングの原動力であるサーバの仮想化技術は、計算資源やストレージ資源を仮想化して独立な利用・運用を可能にすることを目的とするが、このサーバ仮想化技術の成熟により、同様にルー

タ・スイッチのノード資源も仮想化が可能となる。その結果、従来から存在するVLAN、トンネル（カプセル化）、MPLSによるリンク資源の仮想化と成熟しつつあるノード資源の仮想化を組み合わせることで、ネットワークというインフラストラクチャ全体を仮想化することが可能になる。

進化するネットワーク仮想化技術を用いて、我々が構築を進めている新世代ネットワーク情報通信基盤は、広義のネットワークにおける資源の集合を「スライス」として予約し、スライス上に動的に予約、及び、プログラム可能な、仮想的な「リンク」と「ノード」を基本要素とする「仮想ネットワーク」を生成することが可能になる。

3.2 進化するネットワーク仮想化技術の目的

新世代のネットワーク仮想化技術の目的は、資源を分離し、お互いに干渉することのない「スライス」を構成し、そのスライス上に新しい機能を自由に創造可能とすることで、要求の異なる複数のネットワークアーキテクチャやサービスを同時収容する基盤（メタアーキテクチャ）を実現することである。「メタ」という言葉は上位概念を表すが、複数のネットワークアーキテクチャを束ねて包含するという意味でメタアーキテクチャという言葉を使う。

このように、情報爆発時代における多種多様のネットワークサービスやネットワークアーキテクチャを複数同時に収容するためには、従来のような画一的な通信基盤ではなく、進化するネットワーク仮想化技術を用いて、独立した資源の集合であるスライスを構成し、自由にプログラム可能な仮想ネットワークを構築することが必要である。

3.3 ネットワーク仮想化情報通信基盤の要件

我々は、新世代におけるネットワーク仮想化を実現する情報通信基盤は以下の特徴を必須とすると考える。

- (1) ネットワークを構成する資源を体系的に記述し、ユーザやサービスの要求を形式的に指定し割当て可能とする、資源の抽象化（Abstraction）を実現する。
- (2) ネットワークを構成する資源を、ネットワークの任意の範囲に対し、独立分離してスライスに割り当て、パフォーマンス、セキュリティ、名前空間などがスライス間で相互に干渉しない、資源の独立分離（Isolation）を実現する。
- (3) ネットワークを構成する資源を必要に応じて動的に確保・解放する、資源の融通性（Elasticity）

■ 用語解説

GPGPU General-Purpose computation on Graphics Processing Unitsの略。画像処理プロセッサを汎用的な演算処理に応用する技術。

GRE-TAP トンネル GREプロトコルを用いて第2層（レイヤ2）の仮想リンクを実現するトンネリング技術。

JGN2plus NICTが運用している新世代ネットワーク研究開発のためのテストベッドネットワーク。

PlanetLab 米国で2002年に開始した広域のネットワーク実証実験を可能とする世界規模のテストベッド。

CoreLab 我が国で2008年に開始した広域のネットワーク実証実験を可能とするPlanetLabを拡張した我が国独自のテストベッド。

GICTF グローバルクラウド基盤連携技術フォーラム。

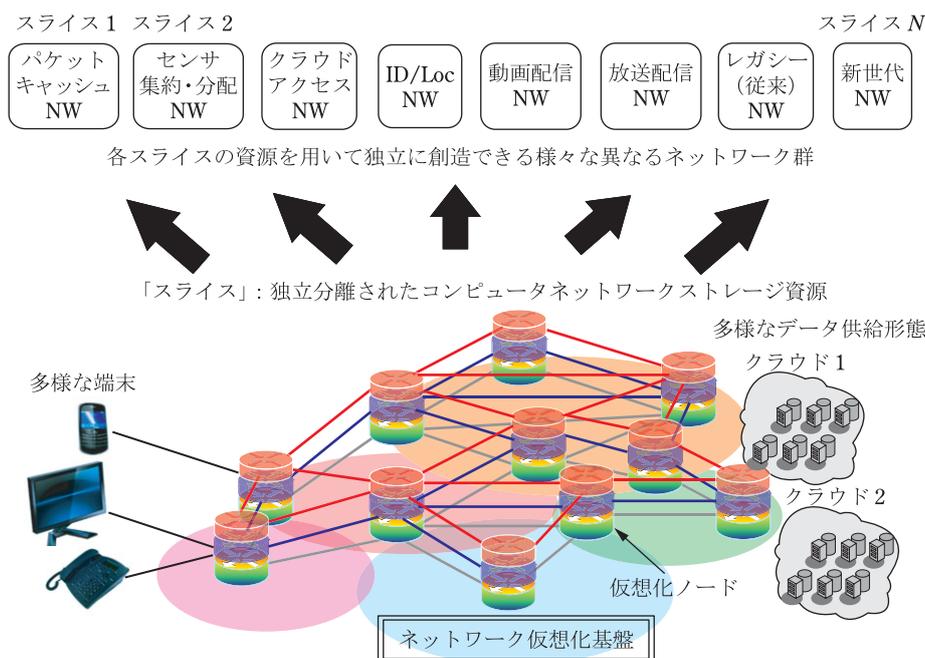


図1 ネットワーク仮想化スライス概念図 新時代に普遍的に存在する多様なネットワークは、ルータ・スイッチの進化形である「仮想化ノード」で構成するネットワーク仮想化基盤で収容される。各ネットワークは、独立分離されたコンピュータネットワークストレージ資源である「スライス」と呼ばれる資源単位上実装される。

を実現する。

- (4) ネットワークにおける計算資源やストレージ資源を駆使し、新しいプロトコルの解釈やデータ加工処理などのプログラム性 (Programmability) を実現する。

- (5) 各種資源の利用を行うための認証・認可・管理の仕組みを備える (Authentication, Authorization and Accounting)。

3.4 ネットワーク仮想化がもたらす利便性

ネットワーク仮想化により実現される新世代の情報社会基盤は、これまで実現できなかった以下のような利便性が情報社会に新たに創出されると期待される。

- (1) アプリケーションごとやユーザごとに独自の QoS (Quality of Service) やネットワーク内処理、セキュリティやプライバシーを設定可能な独立したネットワークが利用可能となる。
- (2) 新たなネットワークサービス利用シーンに必要なネットワーク機能を自由に創ることが可能となり、データ通信の効率性・堅ろう性を高度化することが可能になる。

- (3) 従来のネットワークと複数の新規のネットワークの仕組みを同時に運用することが可能になり、ユーザがネットワークを自由に選択することが可能になる。これにより、従来のネットワークから新しいネットワークへシームレスに持続進化が可能になる。

- (4) 複数の新しいネットワーク実験を同一の物理的なインフラを用いて遂行可能となり、コスト効率の高いテストベッドを構築可能な基盤技術として有用である。

ネットワーク仮想化技術インフラ提供者、ネットワークサービス提供者、ユーザなどステークホルダー間の関係を見直し、新しいビジネスモデルを創り出すことが可能となる。

4. 我が国のネットワーク仮想化基盤技術研究

4.1 設計原理

現在、我々は「仮想化ノード」の研究開発の産官学の共同研究^(注1)において、短期的にはテストベッド構築、長期的にはメタアーキテクチャの実装を目指し、「ネットワーク仮想化基盤」のプロトタイプ構築を行っている

(注1) 我が国では、2010年、東大、NICT、NTT、NEC、日立、富士通研の6者により「仮想化ノード」の産官学共同研究を開始した。

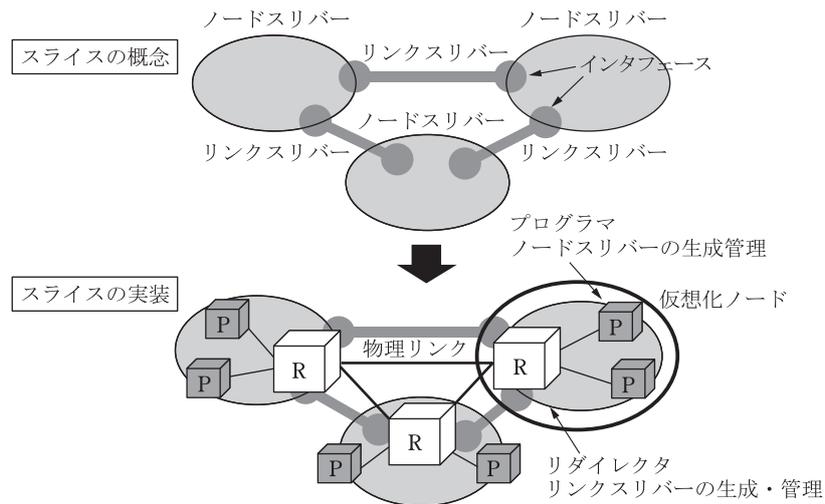


図2 仮想化ノードアーキテクチャ スライスとは、リンクスリバー（仮想リンク）、ノードスリバー（仮想ノード）、インタフェースの集合体として捉えられる。仮想化ノードを構成する物理装置は、リンクスリバーを生成管理するプログラマ、及び、ノードスリバーを生成管理するリダイレクタで構成される。分散して存在する仮想化ノードが構成するネットワーク仮想化基盤上に複数のスライスが生成・管理される。

(図1). このように、新しく進化しつつある「ネットワーク仮想化技術」は、従来のネットワークの限界を打破する新世代のネットワークの重要な根幹をなす先進的研究分野として発展しつつある。

我々はプロトタイプ構築の際の設計原理として、仮想化ネットワーク基盤そのものが持続進化可能な設計を行っている。例えば、ルータやスイッチにおける転送機能やプログラム環境を提供するコンピュータ資源は仮想化のハードウェアアクセラレーションや処理速度の急速な向上などプロセッサの進化により大きく変化をする可能性がある。また、ネットワーク資源に関しても現在は電氣的なパケット転送が主流であるが、例えば、光パス・光パケットなどによる通信技術が日々研究され、より現実的な実装技術となりつつある。構成要素技術は日進月歩発展を遂げる宿命があり、理想的には、構成要素が独立に進化することを許容するネットワーク仮想化基盤の設計が望ましい。

この設計原理に基づき、我々は、まずスライスの概念を、スライスの接続性とプログラム性を分離し、それぞれリンクスリバー (Link Sliver) とノードスリバー (Node Sliver) に分割する (図2)。また、リンクスリバーを構成する要素技術のリダイレクタ (Redirector) と呼び、ノードスリバーを構成する要素技術をプログラマ (Programmer) と呼ぶ。大雑把にいうと、スライスのトポロジー (骨格) を構成するリンクスリバーを提供する要素技術がリダイレクタであり、スライスにおけるプログラム可能性を提供するノードスリバーを実現する要素技術がプログラマである。リダイレクタは通常はネットワーク機器により構成され、プログラマは一般にコンピュータ機器により構成される。

このように、リダイレクタとプログラマという要素技術を独立に扱うことで、それぞれが独立に進化することを許容するネットワーク仮想化基盤を構築することが可能になる。例えば、リダイレクタは、電氣的なパケットを扱うルータをベースとして実装される場合もあるが、光パスを張る装置の実装もあり得る。同様にプログラマも、インテル x86 アーキテクチャのサーバ上の仮想マシン (一般に Slow Path と呼ばれる) で実装される場合もあれば、ネットワークプロセッサや FPGA (一般に Fast Path と呼ばれる)、または、近年科学技術計算などで多用される GPGPU^(用語)などのコンピュータ資源をネットワーク処理に利用する場合にも対応可能である。

従来ネットワークと比較すると、リダイレクタとプログラマの集合は、現在のルータやスイッチの機器にプログラム機能を追加して高機能ルータに進化させた機器 (ネットワークノード) のように見える。我々は、リダイレクタとプログラマの集合を「仮想化ノード (VNode)」と定義し、複数のスライスを提供するための基盤要素技術の一つと考えている。

4.2 ネットワーク仮想化基盤

我々の目指すネットワーク仮想化基盤には、①インフラ提供者、②サービス提供者、③ユーザという3種類のプレーヤが存在する。従来のネットワークでは、①と②は同じ提供者が兼ねているのが一般的であり、インフラ・サービスの提供者とユーザという2種類のプレーヤしか存在しない。つまりネットワークを提供する側とそれを利用する側である。

しかし、ネットワーク仮想化基盤では、①と②の分離こそが、メタアーキテクチャの実現の鍵となる。つま

り、①のインフラ提供者はスライスを提供し、②のサービス提供者はスライスを利用してネットワークを③のユーザに提供し、③のユーザは②のサービス提供者が提供するネットワークを利用するという構造となる。ここで、一般に②のサービス提供者は複数存在することを想定しているため、①のインフラ提供者は、スライスをそれぞれのサービス提供者に提供することでメタアーキテクチャを実現する。

この①と②の分離は、ビジネス的に見ると、インフラ提供者＝サービス提供者という制約が除去され、ネットワーク利用技術の市場開拓につながる可能性がある。モバイルネットワークで見られる MVNO や、同じ航空機の機体をインフラとして複数の別の航空会社のサービスを実施するコードシェア便にも同様のビジネスモデルを見ることができる。

前述の3種類のプレーヤとそのやり取りを管理するため、ネットワーク仮想化基盤には、「ドメインコントローラ (Domain Controller)」と呼ばれる仮想化管理機構が存在する。ドメインコントローラは、①のインフラ提供者が制御を行うための機構であり、一般に、インフラの境界 (ドメイン) に一つ存在する。現在のネットワークでは、インターネットは、ドメインで管理される自律システム (AS) が相互に接続することで「ネットワークのネットワーク」を形成している。同様に、複数ドメインにまたがって形成されるスライスは、複数のドメインコントローラの相互接続により実現される。

ドメインコントローラは、②のサービス提供者からスライスの設計図とともにその作成の要求を受け、ドメイン内の全ての仮想化ノードにスライス構築・廃止を统一的に指示する。ここで、仮想化ノードには、リダイレクタとプログラマを管理する仮想化ノードマネージャが存在する。ドメインコントローラからのスライス構築・廃止の指示は、仮想化ノードマネージャに渡り、スライスのトポロジー構築がリダイレクタに指示され、サービス提供者が用意したプログラムがプログラマにロードされるため、スライスの中に所望のネットワーク機能が実装される。スライスが構築されると、サービス提供者が用意したネットワーク機能は稼働を始め、③のユーザが自由に使用可能となる。スライスの廃止も同様の手順で実行される。

ネットワーク仮想化基盤のもう一つの構成要素にアクセスゲートウェイがある。現在のネットワーク仮想化基盤プロトタイプでは、ユーザがネットワーク仮想化基盤上の複数のスライスにアクセスするための、①認証機構、②スライス選択機能、③データ生成場所とスライスとの中継機能が必要であり、それらをアクセスゲートウェイで実施する。ユーザは、ドメインコントローラに登録を行った後、アクセスゲートウェイで認証を受け、各スライスにアクセスする。アクセスゲートウェイの設

置場所は、現在のネットワーク仮想化基盤プロトタイプでは規定されていない。しかし、メタアーキテクチャのコンテキストではユーザの端末やデータの生成場所との接点に位置し、テストベッドにおいては一般ユーザを収容して実証実験に参加してもらうためテストベッドと従来ネットワークの境界に位置するなどコンテキストにより多様な場合があり得る。また、ユーザとの接点だけではなく、スライス間の相互接続や従来ネットワークとの接続などあらゆる種類のゲートウェイが必要であると考えられ、今後アクセスゲートウェイの定義とプロトタイプ実装は進化すると考えられる。

4.3 プロトタイプ実装

「仮想化ノード」研究開発の共同研究では、仮想化ノード、ネットワーク仮想化基盤管理機構、アクセスゲートウェイから成る持続進化可能なネットワーク仮想化基盤のアーキテクチャを詳細に設計し、そのプロトタイプ実装を行っている。仮想化ノードの内部構造を簡単に述べると、リダイレクタ部分は、商用ルータをベースに独自開発したアドオンカードにリダイレクタ拡張機能を実装し、プログラマ部分は、PC サーバ上の仮想マシンとマルチコアネットワークプロセッサ上のコンピュータ資源を組み合わせて利用できるように構成されている。

現在のプロトタイプでは、リダイレクタはリンクスリバーを GRE-TAP トンネル^(用語)で実装している。これは、現在のプロトタイプが JGN2plus^(用語)上で試験運用を予定しているためだけであり、アーキテクチャとしては、リンクスリバーの実装は VLAN や光パスによる波長も利用可能であり、GRE-TAP トンネルに限定されるものではない。

重要なのは、我々の設計原理によりリンクスリバーの実装は今後進化し多様化するが、ユーザ端末からアクセスゲートウェイ、そして複数の仮想化ノードを経て、再び、異なるユーザ端末に至るまで、スライスの中では、フォーマットを規定しない「フォーマットフリー (Format Free, Any Format)」のデータが流れることを保証することである。つまり、我々のネットワーク仮想化基盤では、エンドツーエンドで「フォーマットフリー」なデータ通信が実現可能である。このとき、仮想化ノードのノードスリバーにおいてはデータをどちらのリンクスリバーに分岐するべきかという判断をしなくてはならないが、これはノードスリバーに実装されたプログラムがデータの転送判断を行うために可能となるわけである。従来ネットワークに例えれば、ルータでは、到達したデータパケットはフォワーディングテーブルを基にインタフェースに分岐して転送するべくルーティングプログラムが実装されているが、ネットワーク仮想化基盤では、そのプログラムが自由に設定できるため、従来とは

全く異なるネットワークがスライス内に実装可能である。

ここで、我々のアーキテクチャは、必ずしもパケットによるデータ転送を規定しているわけではないことに注意されたい。例えば、リダイレクタが光パスを実装する場合は、サーキットスイッチのスライスを構築することも可能であるからである。

このように、我々の考えるネットワーク仮想化基盤は、リンクリバーを流れるエンドツーエンドのフォーマットフリーなデータ通信とノードリバーのプログラムによる分岐により、従来のネットワークと異なる複数の多様なネットワークの仕組みやプロトコルを収容するメタアーキテクチャを実現する。

これまでの研究開発における仮想化ノードのプロトタイプは、10 Gbit/s の転送能力を有する 4 台、アクセスゲートウェイ 7 台、ユーザ端末 15 台を、2010 年 9 月に JGN2plus の実験網に展開し実証実験を展開している。2011 年には仮想化ノードを更に数台追加してより規模の大きいインフラへと進化させ、今後、更にスケラブルな実証実験を遂行する予定である。

4.4 今後の課題

我が国のネットワークテストベッドである JGN2plus 上に展開し、ユーザに試用してもらう試験的な運用を開始し、テストベッドとしてのユーザビリティの検証と機能拡充を行う。JGN2plus に展開するネットワーク仮想化基盤は、ユーザに使ってもらうためのテストベッドであると同時に、それ自体のテストベッド機能を確認するためのテストベッドでもある。

まず、今後の急務は、ネットワーク仮想化基盤のアーキテクチャの拡張とプロトタイプの完成度を高めることである。特に、最新の通信技術やプロセッサ技術を構成要素として取り込むことが重要である。特に我が国が誇る無線通信技術、光通信技術、また、マルチコアプロセッサ、GPGPU などの新しいコンピュータ資源などの分野における共同研究のパートナーの優れたエンジニアリングやキャリアの視点をネットワーク仮想化基盤に生かして国際競争力を養うことが大切である。

ネットワーク仮想化基盤は、長期的な視点で見ると、テストベッドだけではなくメタアーキテクチャを実現する新しいネットワークのインフラである。そのため、新しいインターネットを形成するためのメタアーキテクチャのコア技術として研究開発を進める必要がある。世界各国で、ネットワーク仮想化技術の研究開発、特に、ネットワーク仮想化基盤の研究開発が進んでいる。その

ような研究コミュニティで日本発のアーキテクチャを周知し、また、国際連携のための相互接続 (Federation) を進めなくてはならない。

ネットワーク仮想化基盤の研究開発は、基盤技術を中心とした研究開発に焦点を当てがちであるが、我々が、PlanetLab^(用語)、⁽⁷⁾、⁽⁸⁾ や CoreLab^(用語)、⁽⁹⁾、⁽¹⁰⁾ で培った経験によると、基盤技術の研究開発だけを進めるだけでは不十分である。基盤技術と同時に、基盤の上で実装される個々のネットワークアーキテクチャとともに、両輪での研究開発を進めて初めて成熟した基盤技術となる。特に、GICTF^(用語) で進められるクラウド間連携のためのネットワークには、新しい機能をネットワーク内部に付加することで高機能化、高信頼化を図ることが議論されている。またコンテンツ指向ネットワーク、キャッシュ指向ネットワークなど新たなネットワークの仕組みの研究も盛んに行われつつある。

最後に、進化するネットワーク仮想化の研究成果により、将来的に、利用者にとって使いやすく安全に利用できるネットワークを実現し、また、我が国産業界にとっても新たなネットワーク利用技術開拓による新しい市場創出につなげていく必要がある。

文 献

- (1) National Research Council, Looking Over the Fence at Networks, National Academy Press, Washington D.C., 2001.
- (2) A. Nakao, "Network virtualization as foundation for enabling new network architectures and applications," IEICE Trans. Commun., vol. E93-B, no. 3, pp. 454-457, March 2010.
- (3) GENI: Global Environment for Network Innovations, <http://www.geni.net>
- (4) The OpenFlow Switch Consortium, <http://www.openflowswitch.org/>
- (5) FIND, <http://www.nets-find.net/>
- (6) FP7, http://cordis.europa.eu/fp7/home_en.html
- (7) PlanetLab, <http://www.planet-lab.org>
- (8) L. Peterson, T. Anderson, D. Culler, and T. Roscoe, "A blueprint for introducing disruptive technology into the Internet," Proceedings of HotNets-I Workshop, pp. 59-64, 2002.
- (9) CoreLab, <http://www.corelab.jp>
- (10) A. Nakao, R. Ozaki, and Y. Nishida, "CoreLab: An emerging network testbed employing hosted virtual machine monitor," Proceedings of ACM ROADS Workshop, 2008.

(平成 23 年 1 月 12 日受付 平成 23 年 2 月 4 日最終受付)



なかお ありひろ
中尾 彰宏 (正員)

1991 東大・理・物理卒。1993 同大学院工学系研究科情報工学専攻修士課程了。IBM Texas Austin 研究所、IBM 東京基礎研究所などを経て、米国・プリンストン大大学院情報科学科にて修士号及び Ph.D. 取得。2005 東大大学院情報学環准教授。2007 から NICT 客員研究員兼任。