ルータクラウド・インフラストラクチャに関する研究

代表研究者	鯉 渕 道 紘	国立情報学研究所アーキテクチャ科学研究系准教授
共同研究者	西 宏 章	慶應義塾大学大学院理工学研究科准教授
共同研究者	川 島 英 之	筑波大学大学院システム情報工学研究科専任講師

1 概要

本研究調査では，インターネット・ルータがパケット転送するのみならず，情報の発信・共有・検索・受信に積極的に関わるルータクラウド・インフラストラチャの要素技術を提案，開発した．具体的には，下記(1)(2)などを行い，その成果を統合することで(3)を探求し，本研究調査のまとめを行った．

（１）ルータアーキテクチャの策定（２００９年４月～同年８月）
　　シミュレーション（２００９年９月～同年１２月）
（２）メモリ DB を用いたストリーム処理エンジンの開発（２００９年４月～同年８月）
　　問合せ言語の開発（２００９年９月～同年１２月）
（３）サービスと応用サービス，セキュリティ支援（２００９年９月～同年３月）
先進的インターネットへの適用可能性（２００９年１２月～２０１０年６月）

2 はじめに

近年，インターネット上に存在するあらゆる情報は，マッシュアップなど複数発信源からの情報を組み合わせる手法により，多角的な価値とコンテンツとしての意義を有するようになった．例えば Google や Amazon が次々と生み出す新技術が Web アプリケーション・サービスの高度化に寄与し，新たなビジネスを掘り起こし広げている．

現在，Web アプリケーション・サービスのさらなる高度化のための 1 つの方法として，インターネット・インフラストラクチャであるルータやゲートウェイが取得可能な情報を積極的に活用する，あるいはルータが担うスイッチングをサービスに追加する研究が進められている．

例えば Cisco ISR(Integrated Services Router) 向けに提供されている AXP (Application eXtension Platform) はルータ上で Linux アプリケーションを実行するための API を提供している．また，Active Network では，ネットワークノードがキャッシュを持ち，株式市況やオンラインオークションのサーバーの負荷を軽減するために，トラフィックを解析してコンテンツに応じてパケット処理を最適化することなどが検討されてきた．また，リコンフィギュラブルなハードウエアを用いることでアプリケーション層に及ぶパケット解析を高速に行い，IP ベースではなくコンテンツベースのルーティングを行う研究も行われている．これらの研究はルータが単なる通信基盤に留まらず，次世代のインターネットにおけるサービスの中核になりうることを示している．

我々は，ルータをパケットデータの管理基盤として考え，この管理基盤から有用な情報を正規表現により抽出し，その情報をルーティングや新しいサービスの提供に生かすルータを提案してきた．

本研究調査ではこのルータ群によるクラウドシステムを構築することにより，新世代インターネットを実現することを目指し，要素技術の開発などを行った．

提案ルータは，図１に示した通り，既存のルータに高速な情報抽出を効率的に行うための正規表現プロセッサとオンメモリデータベースへの高速なデータインサーションを行うハードウェアを加えた構成を持つ．

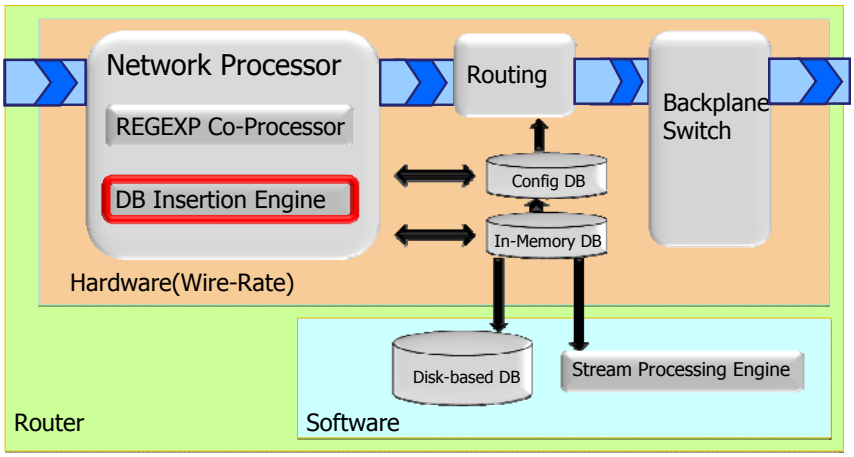


図１：提案ルータの概要

トラフィック情報は，「ある URL にアクセスしたユーザは，他のどんな URL にアクセスするのか」，「ユーザがある URL にどのくらいの時間滞在していたのか」，「その情報がいつ，どこからネットワーク上に現れたのか」といった，検索サービスや人気調査にとって重要な情報を含んでいる．これらの情報は従来のネットワークシステムでは利用が困難であったが，我々が提案を行っているルータ・アーキテクチャではリアルタイムでサービスに必要な情報をネットワークトラフィックから抽出することができる．

以降，３章においてルータ・アーキテクチャとシミュレータ，４章において提案ルータにおけるサービス提供者へのストリーム処理支援，５章において提案ルータの応用範囲およびそのサービス支援について述べる．

３　ルータ・アーキテクチャの策定およびシミュレータ

我々の提案ルータ・アーキテクチャではリアルタイムでサービスに必要な情報をネットワークトラフィックから抽出することができる．よって，サービス指向ルータでは，多くの XML タグを用いたコンテンツベースのルーティングと異なり，TCP パケットのペイロード全体にわたる解析が必要となるため，TCP ストリームの再構築処理が重要となる．

libnids や Snort などで採用されている既存の TCP ストリームの再構築法では，TCP ストリームを構成するすべての TCP パケットを記憶，到着後，組み立てを行っている．そして，アプリケーション側から完全に復元された TCP ストリームに対して情報の検索，抽出処理を行っている．

しかし，これら既存の TCP 再構築法をルータに適用する場合，多数のポートを持ち，かつ，短時間で膨大なストリームデータを扱うため TCP ストリームの再構築におけるメモリ使用量がトラフィック量に応じて大きくなる問題がある．

そこで，サービス指向ルータにおいてパケットから TCP ストリームの再構築情報を抽出し，到着した(TCP)パケット毎に文字列検索を行う部分 TCP 再構築法を提案した．部分 TCP 再構築法は，部分的に TCP パケットを再構築するのみで文字列の抽出情報を判定し，必要なペイロードのみを抽出することでメモリ使用量を抑える．複数パケットに渡る文字列検索を可能とするために部分 TCP 再構築法では各パケットの処理内容をコンテキストとして格納する．

３．１　情報抽出を行うネットワークプロセッサ

ネットワークプロセッサ・ブロックにおける情報選択・抽出は図２に示した通り，以下の５つの処理エンジンにより行われる．

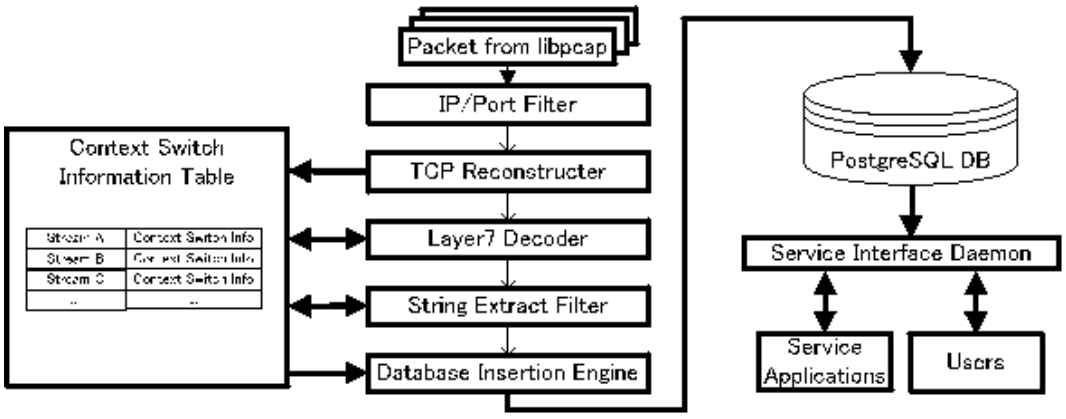


図２：情報選択，抽出機構

(a) IP/ポートフィルタ

ルータの各ポートを通過するパケットは，ネットワークプロセッサ内でまずヘッダ解析される．そしてIP/ポートフィルタ内において IP/Port 情報から解析するパケットを限定する．
具体的にはパケットヘッダ情報から抽出保存が必要ないと判断できるパケットを廃棄する．

(b) TCP 再構築

IP/ポートフィルタを通過したパケットは TCP ストリームへ再構築される．

(c) レイヤ 7 デコーダ

アプリケーションプロトコルのデコードを行う．TCP ストリームからレイヤ 7 デコーダによって，HTTP/1.1 や MIME エンコード等のアプリケーションプロトコルがデコードされる．

(d) 文字抽出フィルタ

レイヤ 7 ペイロードから文字抽出フィルタによって文字列探索が行われる．

(e) データベース挿入エンジン

抽出された文字列をメモリデータベースへと保存する．

3.2 部分 TCP 再構築法

本章では，サービス指向ルータにおける文字列検索が可能な部分 TCP 再構築法を述べる[3]．

3.2.1 既存の TCP 再構築法の問題点

Libnids や Snort などの代表的な既存の TCP 再構築法の実装は，ストリーム全体を保存するパケットバッファが必要となる．
TCP 再構築エンジンとして libnids，および Snort Stream5 の実装をサービス指向ルータへ適用した場合の具体的な挙動を図３に示す．

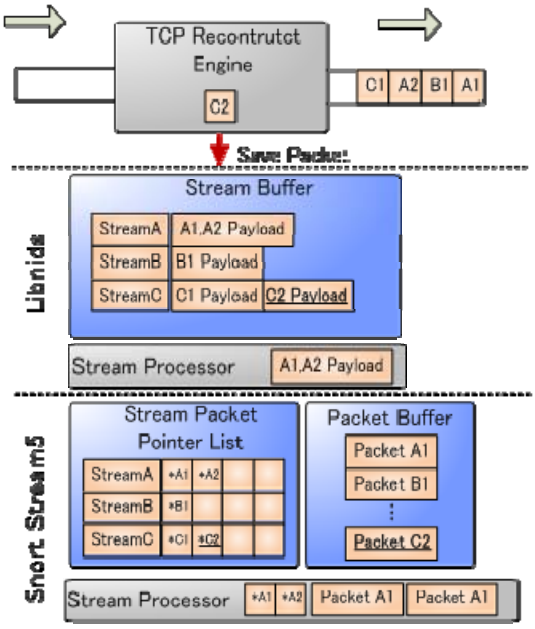


図３：既存の TCP ストリーム再構築法

図３では，TCP ストリーム(ストリーム A,B,C)を構成する TCP パケットが A1,B1,A2,C1 の順に到着し，ストリーム構築が終了したストリーム A に対して情報抽出処理を行っていることを表している．

libnids の実装の場合，ストリームバッファと呼ばれるメモリ領域を用意し，ストリーム毎の TCP パケットのペイロードを書き込んでいく．また，同様の処理を Snort Stream5 で行う場合，メモリ領域に，ストリームを構成する TCP パケットへのポインタを格納するストリームパケットポインタリストと， TCP パケットを格納するパケットバッファを用いる．そして，ストリームプロセッサ部は，そのポインタと TCP パケットの両方を読み込み，情報抽出を行う．

したがって，これらを多数のコネクションが同時に通信を行うルータにおける情報取得手法として採用した場合，各ポートは多数のストリーム全体が揃うまでペイロードを格納する必要がある．

これはメモリ使用量の点で以下の 2 点の無駄を生む．

(a) 非選択 TCP パケットの構築の無駄

例えば IDS に関するパケットを破棄する問合せ集合 Qd を n 個の IP パケットから構成される TCP パケット t に適用することを考える．i(1<= i <= n)番目の IP パケット ip が IDS の特徴を有する場合には，ip 到着時に t は破棄可能であることが Qd の条件から判明する．しかし，既存の TCP 再構築の実装では n 番目の IP パケットが到着するまで t を破棄できない．このとき，j (i+1<= j <= n) 番目の IP パケットを保持するメモリ空間は無駄になる．

(b) TCP パケット中の非抽出ペイロードの構築の無駄

例えばペイロードの一部のみを抽出する問合せ集合 Qe を考える．String Extract Filter において Qe で示されるペイロードの一部が抽出される．一方，残りの非抽出ペイロード部を保持するメモリ空間は無駄となる．

3.2.2 部分 TCP 再構築法の提案

ここでは，メモリ量を削減するために，ストリームを構成するパケットが到着した時点で文字列検索の解

析を開始し，解析の途中状態を保存する部分 TCP 再構築法を提案する．

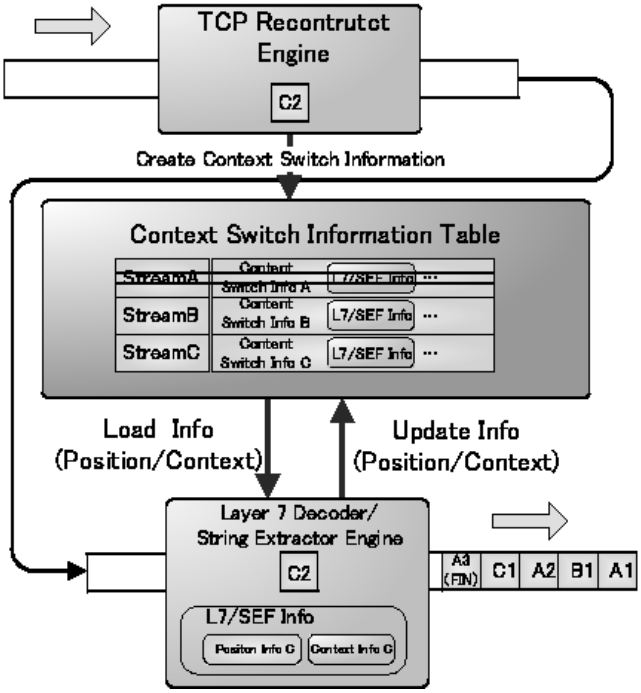


図４：部分 TCP 再構築法

部分 TCP 再構築法では，図４の TCP 再構築において，各 TCP パケットに TCP ストリームの再構築情報を付加する．TCP パケットに SYN フラグがあった場合，3 ウェイハンドシェークが行われることを確認して新しいストリームの開始を検出し，新しいストリームのためのエントリを Context Switch Information Table に作成する．

部分 TCP 再構築法では TCP ストリームが完成する前に TCP パケット単位で選択・抽出処理を行う．それゆえ，libnids，Snort Stream 5 の実装のように TCP ストリームが終了するまで処理を待つ必要がない．

よって，対象とする抽出パターンが多い場合は途中で保存するコンテキストのサイズが大きくなるが，既存の実装と異なり，メモリ使用量が１つのストリームを構成するパケット数や並列に処理中のストリーム数に線形に増加することを抑制することができる．

部分 TCP 再構築法は，(1)対象となる文字列を含まない TCP パケットはそのまま破棄する点，および，(2)文字抽出フィルタにおいて TCP パケット中の非抽出ペイロードの構築を行わない点から，3.2.1 節で指摘した libnids，Snort Stream 5 の TCP 再構築法における問題点を解決することができる．

部分 TCP 再構築法はすでにシミュレータ SRIM 上[3]において実装済みである．さらに本研究調査では，この他に情報抽出を効率良く処理する正規表現エンジン[1]，正規表現エンジンにより抽出した情報をメモリ上に格納するデータベースインサーション機構[2]についても研究開発を行った．

４　ストリーム処理支援

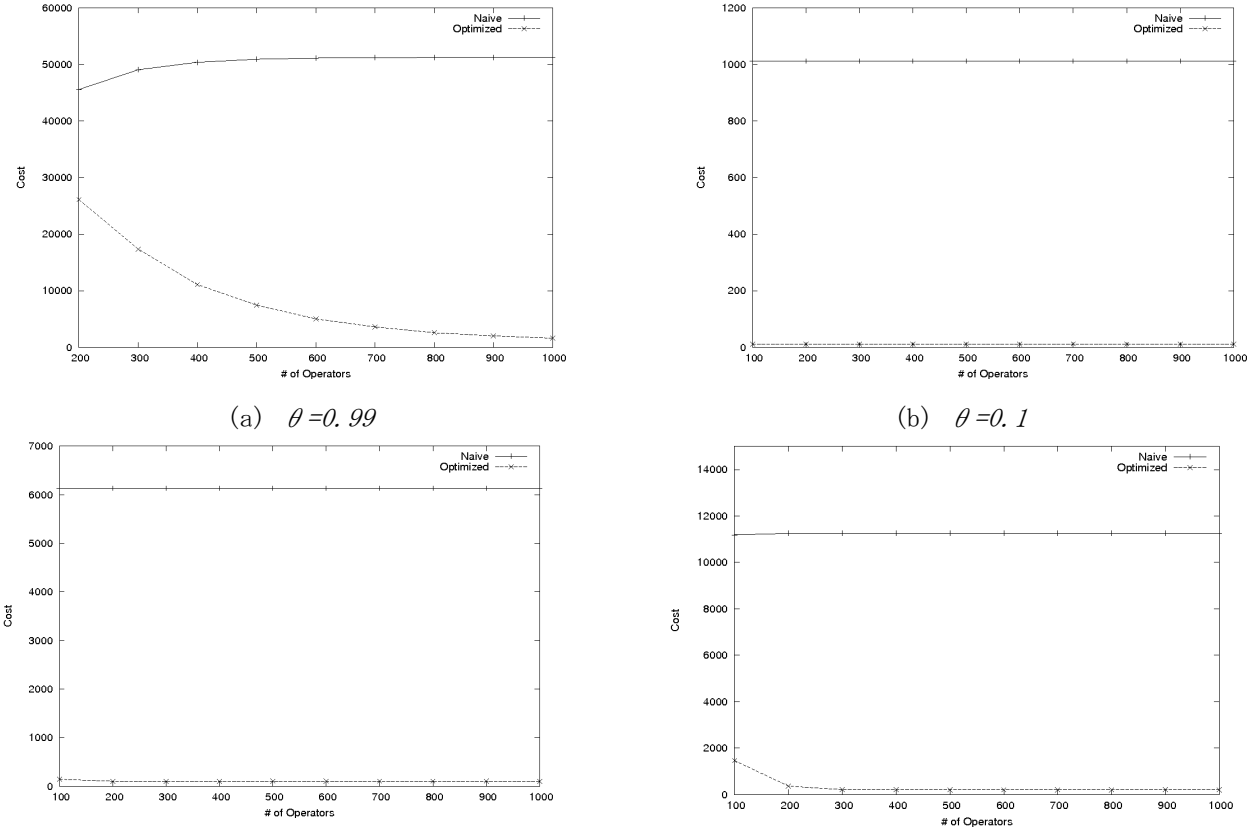
４．１　問い合わせ最適化技法

ユーザ，サービス提供者へ情報抽出に関する高度なプログラミング能力を与えるため，提案ルータ上にストリーム処理エンジン(SPE) を搭載することを提案した[4]．SPE は，SQLライクな宣言的問合せ言語をユーザに提供し，ユーザに与えられた問合せをメモリ上で高速に処理するシステムである．多くのSPE が提供する主たる演算は関係代数演算である．一方，関係代数演算以外の高度な演算子を提供するSPE も現れ始めている．その例には，データマイニング演算を提供するSPADE，信号処理演算を提供するWaveScope，複合イベント処理(CEP) を提供するSASE+，そしてベイジアンネットワーク機能を有するモデルがある．この演算子高度化の流れにおいて，本研究調査ではパケットペイロード処理に頻繁に用いられる正規表現処理を有する選択演算を対象とする．巨大なテキストに対する正規表現処理は負荷が高いため，正規表現処理を有する選択演算は負荷が高い可能性がある．それゆえ関係演算の中で高負荷であることが知られている結合処理と同

等以上の負荷を有する可能性がある．通常の選択演算は低負荷であるため，問合せ最適化処理においては結合処理よりも先に実施することがヒューリスティクスとして広く知られている．しかし，本研究調査で扱う問題においてはこのヒューリスティクスが成立しない．さらに，選択演算子の処理コストが異なる場合の扱いは従来研究では考えられてこなかったが，コストの異なる選択演算子が混在する場合には，それらの実行順序を考慮することで性能が大幅に改善される可能性がある．そこで本研究調査では，正規表現処理を有する高負荷選択演算(REGEXP 選択演算）を含んだ問合せ最適化技法を提案する．具体的には，選択演算子の述語を考慮した静的最適化技法を提案し，REGEXP 選択演算のコストをパケット到着時にオンラインで乱択を用いて見積ることにより，負荷の軽い演算を推定する動的最適化技法を提案した．

４．２ 評価

本節ではシミュレーションにより静的最適化アルゴリズムの評価結果を示す．低負荷，高負荷な選択演算子をn個有する演算木において，従来手法と提案手法の比較を行った．演算木への入力タプル数は100,000 とし，各選択演算子の選択率θ は0.99 に固定した．入力タプルを演算木が処理するコストを計測した．シミュレーション結果を図3,4,5,6 に示す．図の横軸は演算子数を，縦軸は処理コストを表す．



(a) θ =0.99　　(b) θ =0.1

(c) θ =0.9　　(d) θ =0.95

図５： 評価結果

図５に示した実験結果より，提案手法が従来手法よりも優れた結果を示していることがわかる．また，θ が0.95 以下になるとコストが急速に収束することが見て取れる．すなわちθ に対して性能は敏感だと言える．コスト収束する理由は，低負荷演算子の割合が増加することだと考えられる．従来の問合せ最適化では選択演算の負荷を考慮することはなく，選択演算は単純に演算木の下部へとプッシュダウンするだけであった．しかし，我々が想定する環境においては，そのアプローチは単純すぎることが本実験結果は示唆している．提案手法の従来手法に対する欠点は，演算子の整列を行う必要がある点である．すなわち，提案手法はn 個の演算子があるときに$O(n(log(n)))$ だけの計算量が必要になる．これは問合せが揮発的なRDBMS では問題に

なる可能性があるが，問合せが永続的なSPE では問題にならないと考えられる．

５．サービスと応用サービス，セキュリティ支援

提案ルータにより実現される応用サービス,セキュリティ支援等について,以下の幅広い領域毎に示した．

・防犯，セキュリティ：多様な要件の複数ネットワークを単一基盤で同時収容可能なネットワーク仮想化技術，あるいはユーザ単位のセキュアプライベートネットワークをオンデマンドで瞬時に構築できる動的ネットワークリソース共有技術や，ユーザや状況に応じて設定・変更可能な適応的プライバシー保護技術が求められている．本提案ルータは，ネットワーク仮想化の実現に特殊な機器やソフトウェアを利用する手法と比較し，コンテンツに内包する情報による選別とコンテンツベースのネットワーク仮想化手法を構築することでユーザ単位での構築も可能であり，またその設定・変更もオンデマンドで可能となる．

・事故課題：人や車の位置情報等をコンテキストとして流通させるしくみと必要に応じたプライバシー制御が求められている．コンテキストに含まれる位置情報に基づくルーティングが可能であるだけでなく，ルーティングを切り替えることや，プロトコルの変換，QoS の変換，VLAN の乗換などが制御できるため，プライバシー制御もある程度可能となる．

・文化・生活の多様性：個別の機能はシンプルでかつ高速で動作するが，それらを自在に組み合わせて利用できるようなフレームワークとしてデータ配置の最適化と省電力アクセス技術を新世代ネットワークが持ち合わせることが必要とされている．提案ルータが提供するフレームワークはシンプルかつ拡張性に優れる．HTML5 以降，クライアント側ブラウザがデータベースを持ち，アプリケーションが SQL でデータアクセスする中で，ルータが同様にデータベースを持ち，セキュリティや負荷分散の観点から，データ配置最適化の自由度がさらに向上すると考えられる．提案ルータのパケット解析能力により，物理的に近隣であることだけでなく，精神的，文化的に近傍であることなどを分析することも可能と考えられる．

・メディア融合：複数伝送路から得られるコンテンツをマージして提示する技術また複数コンテンツの違和感のない同期制御技術，容易に情報配信可能なメディア融合プラットフォームの構築が求められている．Web サービスを活性化し利便性を高めているサーバ間通信，いわゆるマッシュアップは，提案ルータにより一層活性することができる．提案ルータが獲得できるデータはユニークで，カバレッジと最新性に優れる．マージや同期制御は各アプリケーションで行い，提案ルータはそれらユニークなデータへアクセスできる API を提供する．

これらのアプリケーションはすべて，(1) 正規表現に基づいて選択・抽出したコンテンツを，(2)データベースによる蓄積・選択・抽出を通して，(3)ルーティングテーブルやその他の様々な情報とともに管理し，(4)ユーザ，アプリケーション側へSQLを拡張したAPIで操作させる柔軟性を提供するという一貫した提案ルータを基盤としたクロスレイヤ技術による．

６．まとめ

本研究調査では，インターネット・ルータがパケット転送するのみならず，情報の発信・共有・検索・受信に積極的に関わるルータクラウド・インフラストラチャを提案，探求することを目的に行った．具体的には，以下の３要素技術などについて研究調査を行った．

（１）ルータアーキテクチャの策定とシミュレーション
（２）メモリ DB を用いたストリーム処理エンジンと問合せ言語の開発
（３）サービスと応用サービス，セキュリティ支援，先進的インターネットへの適用可能性

これらの成果から，より具体化した将来のインターネット像を示すために，平成２２年５月より情報通信研究機構「新世代ネットワーク技術戦略の実現に向けた 萌芽的研究」の公募研究において研究分担者：西を代表として，研究分担者：川島，研究代表者：鯉渕と日立情報通信エンジニアリング（株）を加えた産学連携による開発を行い，簡単なプロトライプを平成２３年３月までに完成させ，成果を継続的に発信し続ける予定である．

【参考文献】

[1] 永冨泰次, 鯉渕道紘, 川島 英之, 西 宏章 ``パケットストリームの正規表現処理を可能とするネットワークプロセッサ'', 情報処理学会研究報告 2009-ARC, Aug 2009
[2] 牧野 友昭, 辻 良繁, 川島 英之, 鯉渕 道紘, 西 宏章, ``サービス指向型ルータにおける高速な書き込み機構の提案'', 電子情報通信学会技術研究報告 CPSY2009-23, pp.79-84, Aug 2009
[3] 石田 慎一 原島 真悟, 川島 英之, 鯉渕 道紘, 西 宏章,``パケ"トデータ管理 盤における抽出処の効率化技法", 電子情報通信学会技術研究報告 CPSY2009-86, pp.309-314, Mar 2010
[4] 川島 英之, 鯉渕 道紘, 西 宏章, ``パケットストリーム処理における正規表現選択演算を含む問合せ最適化", 電子情報通信学会技術研究報告 CPSY2009-87, pp.315-320, Mar 2010

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
パケットストリームの正規表現処理を可能とするネットワークプロセッサ	情報処理学会研究報告 2009-ARC(CD-ROM)	2009 年 8 月
サービス指向型ルータにおける高速な書き込み機構の提案'	電子情報通信学会技術研究報告 CPSY2009-23, pp. 79-84	2009 年 8 月
パケットデータ管理基盤における抽出処の効率化技法	電子情報通信学会技術研究報告 CPSY2009-86, pp. 309-314	2010 年 3 月
パケットストリーム処理における正規表現選択演算を含む問合せ最適化	電子情報通信学会技術研究報告 CPSY2009-87, pp. 315-320	2010 年 3 月