

Realizing Name-based Routing in the Network Layer

黄 恵聖
 大阪大学 大学院情報科学研究科
 情報ネットワーク学専攻
 先進ネットワークアーキテクチャ講座

研究の背景

- ネットワークの利用目的
 - 既知の相手 (who, where) との通信 → 情報 (what) の検索・取得
- ルーティング
 - IP アドレス → 情報・コンテンツ・資源
 - FQDN (ホストの名前) ⇒ Chap. 4
 - 一般化 (資源の名前) ⇒ Chap. 5

問題点

- 多様なコンテンツ・ユーザ数の増加による検索テーブルの巨大化・高速化が必要
 - 検索テーブルのハードウェア構造 ⇒ Chap. 3

FQDN: Fully Qualified Domain Name
 2011.12.5 博士学位論文 公聴会 2

研究の背景

- コンテンツに基づくネットワーク
 - EU の FP7 プロジェクト (2008年~)
 - 4WARD: 特定のホストへのコネクションのみならず、ID が付けられているコンテンツへのアクセスが可能
 - NSF の FIA プロジェクト (2010年~)
 - Named Data Networking (NDN): ユーザは名前でコンテンツを要求し、マッチする情報をルータが転送

FP7: Framework Programme NSF: National Science Foundation FIA: Future Internet Architecture
 2011.12.5 博士学位論文 公聴会 3

研究の目的と概要

ネットワーク層で「名前に基づくルーティング」が実現可能性のあることを明らかにする

- アクセス制御リストが占めるメモリ容量を削減する範囲比較回路と最適化されたプレフィックス展開アルゴリズムを提案・評価
- ホストの名前のルーティング情報を、現存するルータに分散するアルゴリズムを提案・評価
- 資源の名前と、その資源を購読するユーザに関する情報を管理する検索テーブルの提案・評価

2011.12.5 博士学位論文 公聴会 4

博士論文の構成

- Chapter 1 Introduction
- Chapter 2 Related Work
- Chapter 3 A New TCAM Architecture for Managing ACL in Routers
- Chapter 4 Resource Name-based Routing in the Network Layer
- Chapter 5 A New Memory Architecture for Realizing Name Lookup Tables in Resource-centric Networks
- Chapter 6 Conclusion and Future Work

2011.12.5 博士学位論文 公聴会 5

Chapter 3

A New TCAM Architecture for Managing ACL in Routers

1. H. Hwang, S. Ata, K. Yamamoto, K. Inoue, and M. Murata, "A New TCAM Architecture for Managing ACL in Routers," IEICE Transactions on Communications, vol. E93-B, no. 11, pp. 3004-3012, Nov. 2010.
 1. H. Hwang, K. Yamamoto, S. Ata, K. Inoue, and M. Murata, "Minimization of ACL Storage by Adding Minimal Hardware of Range Matching and Logical Gates to TCAM," in Proceedings of the 9th IEEE International Conference on High Performance Switching and Routing (HPSR 2008), pp. 116-122, May 2008.
 1. 黄恵聖, 山本研次, 阿多健彦, 井上一成, 村田正孝, "プレフィックス展開と比較回路の併用によるアクセス制御リストの効率的な管理," 電子情報通信学会技術研究報告 (IN2007-105), vol.107, no.378, pp.37-42, Dec. 2007.
 2. 阿多健彦, 黄恵聖, 山本研次, 井上一成, 村田正孝, "配線ストリマ・転送専用TCAM における効率的なルーティングテーブル管理法," 電子情報通信学会技術研究報告 (NS2007-120), vol.107, no.443, pp.7-12, Jan. 2008

2011.12.5 博士学位論文 公聴会 6

Osaka University Advanced Network Architecture Research Group

本章の背景および目的

- 高速なパケット転送・分類のために使用されている TCAM の問題点: 高い消費電力・チップのコスト
- アクセス制御リスト (access control list; ACL) フィールドの中、ポート番号の範囲表現問題: TCAM 容量の消費量増加

ACL が占める TCAM の容量を削減し、TCAM の消費電力とコストを下げる

TCAM の特性を把握し、名前に基づくルーティングに用いる

TCAM: Ternary Content Addressable Memory

2011.12.5 博士学位論文 公聴会 7

Osaka University Advanced Network Architecture Research Group

提案手法の概要

- 範囲比較回路 (range matching device; RMD)
 - ポート番号の範囲 FROM ~ TO を格納し、入力パケットのポート番号が範囲内にあるのかを判定
- 最適化されたプレフィックス展開 (prefix expansion; PE), PE-MIN
 - AND と NOT ゲートを使用し、従来の PE を改善

PE の例
 $8 \sim 15 \rightarrow 1***$
 (TCAM のメモリセルが 0, 1, * の三つの状態を表現できることを使用)

2011.12.5 博士学位論文 公聴会 8

Osaka University Advanced Network Architecture Research Group

プレフィックス展開アルゴリズム

ポートの範囲指定の例) "5000 ~ 6000"

- PE-OR: 従来のプレフィックス展開

$\begin{matrix} \#5008-5023 & \#5056-5119 & \#5632-5887 & \#5984-5999 \\ \#5000-5007 & \#5024-5055 & \#5120-5631 & \#5888-5951 \\ \#5952-5983 & \#6000 & & \end{matrix}$

$A \vee B \vee C \vee D \vee E \vee F \vee G \vee H \vee I \vee J \Rightarrow 10$ 行

- PE-MIN: 提案方式のプレフィックス展開

$\begin{matrix} 5000: 1001110001000 \\ 6000: 1011101110000 \\ 5120: 1010000000000 \end{matrix}$

$\begin{matrix} \#4992-5119 & \#5120 & \#6000 \\ \#4992-4999 & & \#6000-6015 \\ & & \#6016-6143 \end{matrix}$

$(A \vee B) \wedge (-C \wedge -D \wedge -E) \vee F \Rightarrow 6$ 行

2011.12.5 博士学位論文 公聴会 9

Osaka University Advanced Network Architecture Research Group

TCAM + 論理 NOT/AND 追加回路

論理 NOT (4 Set)
 論理 AND (4 Set)

$(A \vee B) \wedge (-C \wedge -D \wedge -E) \vee F \Rightarrow 6$ 行

- PE-MIN の結果を格納するために、既存の TCAM に NOT と AND の論理ゲートを追加

2011.12.5 博士学位論文 公聴会 10

Osaka University Advanced Network Architecture Research Group

RMD によるエントリ削減効果

- 大阪大学の ACL データベースを TCAM に格納する際に必要なエントリ数 (括弧内は RMD の数)

- プレフィックス展開 (PE) 無し: 3,171,537 (0) → 約500行 (35)
- PE-OR: 1,072 (0) → 約500行 (8)
- PE-MIN: 834 (0) → 約500行 (6)

2011.12.5 博士学位論文 公聴会 11

Osaka University Advanced Network Architecture Research Group

VLSI 上に占める面積の検証

Control Logic: GLUE + ColDrv = 8% 305K Gates

<TCAM VLSI in 90nm Technology> RMD: 580 Gates x 20 RMDs = 11.6K Gates
 Current TCAM: TCAM with RMD = 100 : 100.3

- 既存の TCAM に、RMD を 20個搭載する場合、シリコンオーバーヘッドは 0.3% 増加

2011.12.5 博士学位論文 公聴会 12

Osaka University Advanced Network Architecture Research Group

本章のまとめ

- 範囲比較回路とプレフィックス展開を併用した新しい TCAM を提案し、実運用 ACL を使用し評価
 - ポートの範囲が占める TCAM の行数が 50% 減少 → チップの使用面積を減らす事によって、メモリでの消費電力を半分にすることが可能

TCAM の「*」値を利用した範囲指定:
Chapter 4と5での提案アルゴリズムで使用

2011.12.5 博士学位論文 公聴会 13

Osaka University Advanced Network Architecture Research Group

Chapter 4

Resource Name-based Routing in the Network Layer

2. H. Hwang, S. Ata, and M. Murata, "Resource Name-based Routing in the Network Layer," submitted to Journal of Network and Systems Management (JNSM), June 2011, Revised Sep. 2011.
 2. H. Hwang, S. Ata, and M. Murata, "A Feasibility Evaluation on Name-based Routing," in Proceedings of the 9th IEEE International Workshop on IP Operations and Management (IPOM 2009), pp. 130-142, Oct. 2009.
 3. H. Hwang, S. Ata, and M. Murata, "The Impact of FQDN Database Updates on Name-based Routing Architecture," in Proceedings of the 5th IEEE/IFIP International Workshop on Broadband Convergence Networks (Bcn 2010), pp. 16-21, Apr. 2010.
 4. H. Hwang, S. Ata, and M. Murata, "Frequency-aware Reconstruction of Forwarding Tables in Name-based Routing," in Proceedings of the 5th International Conference on Future Internet Technologies (CFI 2010), pp. 45-50, June 2010.
 3. 黄厚熙, 阿多倫, 村田正幸, "ルータによる名前ベースルーティングの実現に関する一検討," 電子情報通信学会技術研究報告 (IN2008-178), vol.108, no.458, pp.273-278, Mar. 2009.
 4. 黄厚熙, 阿多倫, 村田正幸, "名前ベースルーティングにおける論理・物理トポロジーのマッピング," 電子情報通信学会技術研究報告 (IN2009-167), vol.109, no.449, pp.139-144, Mar. 2010.

2011.12.5 博士学位論文 公聴会 14

Osaka University Advanced Network Architecture Research Group

本章の背景および目的

- ネットワーク層で名前に基づくルーティングを行うメリット
 - 名前を IP に解決する必要性の減少
 - ルータ以外のデバイスで名前を場所に解決する負担の減少
 - ネットワークのオペラビリティの増加

ドメインネームルーティングがハードウェア的に実現可能性のあることを明らかにする

2011.12.5 博士学位論文 公聴会 15

Osaka University Advanced Network Architecture Research Group

提案手法の概要

- FQDN は IP アドレスより長い可変長の文字列
- FQDN の分散アルゴリズムを提案
 - Hierarchical Longest Alphabet Matching (HLAM)
 - プレフィックス展開に着想を得た方法
 - TCAM の特徴を利用
 - Hybrid Distribution (HD)
 - ハッシュ関数による均一な分散の利点を利用した方法
 - 純粋なハッシュ関数を使用するより、閉じたネットワークでのルーティングが可能

2011.12.5 博士学位論文 公聴会 16

Osaka University Advanced Network Architecture Research Group

Hierarchical Longest Alphabet Matching

Content	TCAM cell
Hyphen, digit	01*****
a ~ c	11000**
d ~ g	11001**
h ~ o	1101***
p ~ s	11100**
t ~ w	11101**
x ~ z	1111***

ユニークな 2nd レベルドメイン名のみを格納
Ex) jp.technology.bio
jp.technology.elec
jp.technology.econ
⇒ jp.technology.*のみ格納

2011.12.5 博士学位論文 公聴会 17

Osaka University Advanced Network Architecture Research Group

Hybrid Distribution

FQDN は TLD で分けられ、2nd レベルドメインネームをハッシュ関数をかけて得られた値で区別

TLD: Top Level Domain

2011.12.5 博士学位論文 公聴会 18

Osaka University Advanced Network Architecture Research Group

ルータの使用率と必要な台数

• 閾値 (threshold) : 現存するドメインネームを格納できる最大限
 • 使用率 (utilization) : ルータに格納された比率

$$\text{Utilization} = \frac{\text{number of entries in a router} \times 180 \text{ bits}}{\text{router memory size}}$$
 (router memory size: 1台のルータには18 Mbit TCAM が10個あると仮定)

使用データベース : 現存する約7億個のFQDN
 HLAM: 1,400 台
 HD: 950 台

2011.12.5 博士学位論文 公聴会 19

Osaka University Advanced Network Architecture Research Group

本章のまとめ

- FQDN データベースの分散配置方法を提案
 - Hierarchical Longest Alphabet Matching
 - Hybrid Distribution
- ドメインネームのルーティング情報をルータに格納するために必要なルータの台数は、現存するルータの台数より二桁少ない
→ ハードウェア的に実現可能性がある

2011.12.5 博士学位論文 公聴会 20

Osaka University Advanced Network Architecture Research Group

Chapter 5

A New Memory Architecture for Realizing Name Lookup Tables in Resource-centric Networks

3. H. Hwang, S. Ata, K. Inoue, and M. Murata, "A New Memory Architecture for Realizing Name Lookup Tables in Content-centric Networks," Submitted to Elsevier Computer Networks, Oct. 2011.
 S. H. Hwang, S. Ata, and M. Murata, "Realization of Name Lookup Table in Routers Towards Content-centric Networks," in Proceedings of the 7th International Conference on Network and Service Management (ICNSM 2011), Oct. 2011.
 5. 眞藤聖, 岡多慎吾, 村田正幸, "コンテンツセントリックネットワークを可能にするルータによる名前検索テーブルの実現," 電子情報通信学会技術研究報告 (N2010-149), vol.110, no.449, pp.31-36, Mar. 2011.

2011.12.5 博士学位論文 公聴会 21

Osaka University Advanced Network Architecture Research Group

本章の背景および目的

- ルーティング
 - IP アドレス → 情報・コンテンツ・資源
 - FQDN (ホストの名前)
 - 一般化 (資源の名前)
- 多様なコンテンツとユーザ数の増加

コンテンツとユーザ数の膨大な情報を格納できる検索テーブルの提案・評価

The annual growth is 35.2%
 Data from IDC, NEC, UCSD, NetScout

2011.12.5 博士学位論文 公聴会 22

Osaka University Advanced Network Architecture Research Group

Publish/subscribe (pub/sub)

発行されたコンテンツと、そのコンテンツに興味を持つ購読者の情報をブローカが管理

2011.12.5 博士学位論文 公聴会 23

Osaka University Advanced Network Architecture Research Group

従来のマルチキャスト方式の問題点

Topic	Output interface
⋮	⋮
Apple	fa1 fa2
⋮	⋮

購読者の増加?

- 従来のマルチキャストハードウェアでの問題点
→ メモリの制約
 - 一つのトピックを購読するユーザの数は数万から数十万オーダーであり、zipf 分布に類似
- 対策
 - トピック・購読者情報を格納するメモリの工夫

2011.12.5 博士学位論文 公聴会 24

提案手法の概要

- コンテンツを表す資源名と購読者情報が格納してある検索テーブルの機能をルータで実現
 - マルチキャスト：従来のハードウェアアーキテクチャを利用
 - マルチマッチ：通常は一回の検索に一つの結果しか返さない TCAM を複数回検索

2011.12.5 博士学位論文 公聴会 25

ルックアップテーブルのシナリオ

(a) Active TCAM, passive SRAM
 (b) Passive TCAM, active SRAM
 (c) Passive SRAM, active DRAM

2011.12.5 博士学位論文 公聴会 26

メモリチップのコスト & 遅延

シミュレーションによる評価

- 横軸：購読者情報が格納されるメモリの横ビット
- コスト：使用されたメモリに比例。A > B > C
- 遅延：最もポピュラーな資源名を購読しているユーザ情報を返す時に必要な時間。C > A > B

2011.12.5 博士学位論文 公聴会 27

遅延 & 使用率 with コスト制限

- SRAM/DRAM の横ビット長 (36) とメモリ用の予算 (\$1,000) を固定して評価
- 遅延：C > A > B, 使用率：A > B > C

2011.12.5 博士学位論文 公聴会 28

本章のまとめ

- コンテンツとユーザの膨大な情報を、ルータに格納するための検索テーブルの提案・評価
 - 資源名とユーザ数は一般的に zipf 分布を持つため、データベースの特徴を考慮して検索テーブルを設計する必要がある
 - 高速なメモリを使用しているためコストはシナリオ C より高いが、遅延は 10% である、passive TCAM & active SRAM (シナリオ B) が適切

2011.12.5 博士学位論文 公聴会 29

本研究のまとめ

- ネットワーク層で「名前に基づくルーティング」が実現可能性のあることを明らかにした
 - 範囲比較回路とプレフィックス展開アルゴリズムを併用し、TCAM で範囲が占める容量が約 50% 減少
 - ホスト名のルーティング情報は現存するルータの台数より二桁少ない数で格納が可能
 - 資源の名前と受信者に関する情報を管理する検索テーブルの構造は遅延とコストが低い、passive TCAM & active SRAM のシナリオが適切
- 今後の課題
 - モビリティ：ローミングホストのサポート
 - ルーティングパス：名前で区別される資源までの最適パス
 - セキュリティ：信頼性のある名前・資源・発行者

2011.12.5 博士学位論文 公聴会 30