

5T-07 IP アドレス検索処理の超高速パイプライン実現法

森川大智 岩田誠 寺田浩詔
高知工科大学情報システム工学科

1. はじめに

近年、情報ネットワーク内のデータ通信量は著しく増加しており、ネットワークの高速大容量化が急務となっている。数 T bps を超える DWDM などの伝送技術が着々と進展する一方で、これらの伝送路を多数収容可能なルータの実現が切望されている。

本稿では、ルータにおけるパケット処理のボトルネックの一つである、IP アドレス検索処理を、パイプライン処理能力に優れたデータ駆動プロセッサコア上で超高速に実現する方式を提案する。

2. Pipelined LC-Trie アルゴリズムの実現法

IP ルーティング処理は、ルーティングテーブルを対象にして、入力パケットの宛先アドレスに一致する最長のエントリを検索する複雑な Longest Prefix Match (LPM) 検索を必要とする。これまでに様々な LPM 検索手法が提案されているが、多くはアルゴリズムの提案であり、ハードウェアを含めた超高速実現法の提案は少ない。

本研究では、メモリアクセス回数の削減、およびパイプライン並列による性能向上を達成するために、平均メモリアクセス回数の少ない LC(Level Compressed)-Trie 法 [2] を採用した。LC-Trie 構造はルーティングテーブル中のエントリを Trie 木構造に写像して深さ方向に圧縮し、平均メモリアクセス回数を極小化する手法である。この LC-Trie 構造を対象に LPM 検索を行う処理の流れを図 1 に示す。

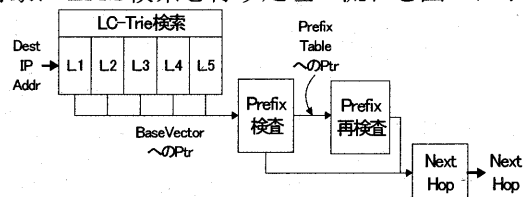


図1 LPM 検索のパイプライン処理構造

図1を現行のデータ駆動プロセッサ DDMP [1]でソフトウェア的に実現すると、命令セットが単純なため、プログラム全体の命令数およびクリティカルパス長が大きくなり、LPM 検索を効率良く実現できない。そこで、(1) 命令の複合化によりプログラム中の命令数を極小化し、検索レートの向上と同時に検索時間の短縮を図る。さらに、(2) 異なるプロトコルにより宛先アドレスのデータ長が異なっても

A Super-Pipelined Implementation of IP Address Lookup
Daichi Morikawa, Makoto Iwata, Hiroaki Terada
Kochi University of Technology
185 Tosayamada, Kami, Kochi 782-8502, Japan

柔軟に対応できるように、入出力データを IP アドレスと Trie 検索に必要なデータに分けて扱う。

これらの要件を満たす複合命令の導入によって、図1の各パイプライン処理ブロックを次のように実現する。

(a) LC-Trie 検索: 宛先アドレスを入力し、LC-Trie 構造を5段（図1: 段数を L で表記）のパイプラインで検索する。IP アドレスおよび検索結果（BaseVector への Ptr.）は Prefix 検査へ移動する。

これらの1段分の LC-Trie 検索のために図2の複合命令を用いる。

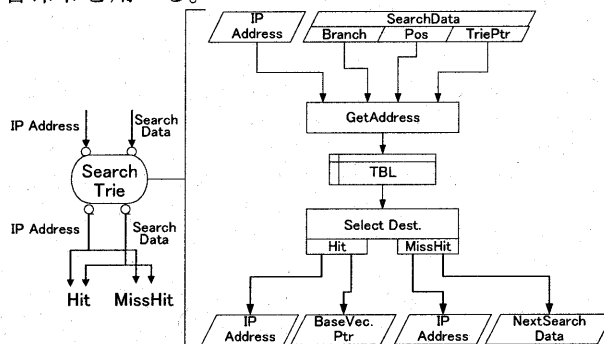


図2 SearchTrie 命令と動作内容

(b) Prefix 検査: Trie の検索において得られた Prefix が正しいか検査する。正しいければ NextHop 参照に移り、誤りならば Prefix 再検査に移る。

(c) Prefix 再検査: PrefixTable への Ptr. が指す Prefix との再検査によって得られる NextHopTable への Ptr. をもって NextHop 参照に移る。

(d) NextHop 参照: Prefix 検査もしくは再検査で得た NextHopTable への Ptr. で NextHop を参照する。

ステップ b と c は同一処理内容であり、ステップ d で参照するテーブルは小さいため、これらを図3に示す複合命令で実現する。

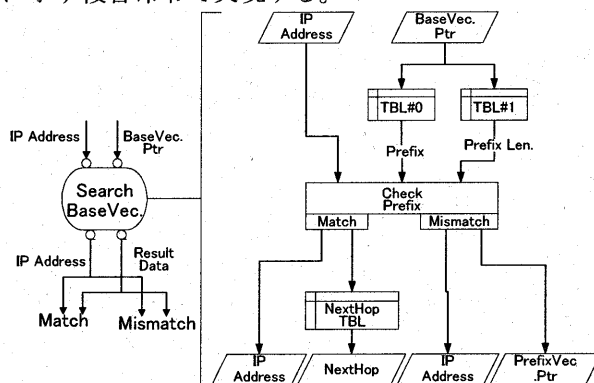
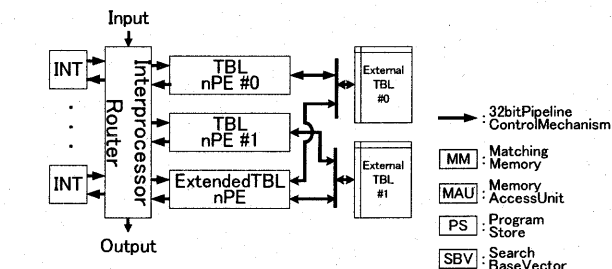


図3 SearchBaseVec. 命令と動作内容

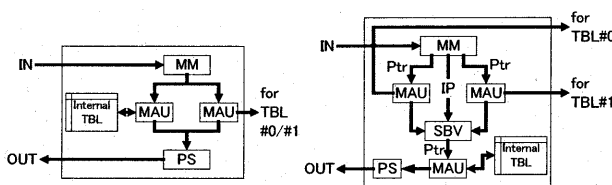
3. IP Lookup 用プロセッサコアの構成法

2章に述べた実現法の実装に際しては、専用ハードウェア実現法の欠点である柔軟性の欠如、汎用プロセッサによる実現法の欠点である低性能の両者を克服するために、次の要件を満足するプロセッサコアを構成する必要がある。

- (a) 必要最低限のハードウェア追加で複合命令を実現でき、既存の命令セットも実行可能なこと。
- (b) メモリアクセスの並列パイプライン化によって、メモリ参照遅延を隠蔽できること。
- (c) ルーティングテーブル内の Prefix 分布に応じて、最適なプログラム、およびメモリ割当てを柔軟に設定可能なこと。



(a)マルチプロセッサコア構成



(b)TBLnPE内部構成

(c)ExtendedTBLnPE内部構成

図4 データ駆動型マルチプロセッサ構成

図4に示すデータ駆動型マルチプロセッサ構成において、オンチップメモリ(内部テーブル)用とオフチップメモリ(外部テーブル)用のアクセス I/F を有する既存プロセッサコアであるテーブル参照専用ナノ PE (以下 TBL) で LC-Trie 検索を行う。さらに、拡張テーブル参照ナノ PE (以下 ETBL) で Prefix 検査、もしくは Prefix 再検査をし、NextHop を出力する。

TBL (図4(b))には、LC-Trie 検索用複合命令 SearchTrie を追加実装することによって、検索処理を実現する。SearchTrie 命令としては、内部テーブルアクセス用と外部テーブルアクセス用を実装する。これによって、LC-Trie 構造中の各レベル毎の prefix 分布数に応じて、内部テーブルもしくは外部テーブルに格納された Trie データを、プログラムによって選択的にアクセスする。

ETBL (図4(c))は2つの外部テーブルへのアクセス I/F と1つの内部テーブルを有する拡張プロセッサコアである。ETBL では Prefix 検査に必要なデータを外部テーブルから取得する。このとき、2つの外部テーブルに同時並行にアクセスし、1回のメモリアクセスと同等の時間で2ワードを参照する。

このようなプロセッサコアを用いたマルチプロセッサを構成する際、複数の TBL を実装することによって、オフチップメモリへのアクセスレートで、LPM 検索レートを達成することが可能になる。これは、各 TBL の負荷を分散すれば、ETBL でのメモリアクセスレートが全体の性能を制約するためである。

本マルチプロセッサ構成では、オンチップメモリサイズが不足する場合には、性能劣化を引き起こさずにマルチチップ化することも可能であるため、大規模なコアルータにも適用可能である。

4. 考察

上述のプロセッサコアの実装によって、大規模なルーティングテーブルに対しても高い LPM 検索レートを達成可能なことを確認するために、北米インターネットバックボーンノード MacEast のルーティングテーブル(IPv4)を対象に性能を見積もった。

現行 DDMP と同等の同等の集積化技術を用いると仮定して、プロセッサ内の処理対象データのビット長 32bit、自己タイミング型パイプラインの動作速度 150M データ/秒、内部テーブル容量 TBL に 128KB、ETBL に 1KB、外部テーブルへのアクセスレート 66M ワード/秒とした時の見積もり結果を表1に示す。

表1 平均検索レートと必要内部メモリ容量

エントリ(KEntry)	55	110	550
平均検索レート	51.06	43.28	36.68
内部メモリ容量	242	260	1572

平均検索レート: Mlookup/sec, 内部メモリ容量 Kbyte

この結果より、本実現法を用いた場合、ルーティングテーブルの増加に伴い、必要な内部メモリ容量は増加するが、最小 IP パケット長 47 バイトの場合でも、18Gbps~13Gbps の性能を達成できる可能性があることが確認された。

5. むすび

本稿では、超高速 IP アドレス検索のためのデータ駆動プロセッサの実現方法を提案した。本方式は、並列パイプライン処理により IP アドレス検索のボトルネックとなるメモリアクセス遅延を隠蔽可能なため外付けメモリのバンド幅と等しいレートで最大検索レートが達成可能である。今後、詳細設計ならびに回路レベルの評価を実施する予定である。

参考文献

- [1] H.Terada, et al., "DDMP;s: self-timed super-pipelined data-driven multimedia processors", *Proc. of the IEEE*, vol.87, no.2, pp.282-296, Feb. 1999.
- [2] S.Nilsson and G.Larlsson, "IP-address lookup using LC-Tries", *IEEE J. on Selected Areas in Comm.*, vol.17, no.6 pp.1083-1092, Jun. 1999.