Poptrie: A Compressed Trie with Population Count for Fast and Scalable Software IP Routing Table Lookup

Hirochika Asai
The University of Tokyo
panda@hongo.wide.ad.jp

Yasuhiro Ohara
NTT Communications Corporation
yasuhiro.ohara@ntt.com

ABSTRACT

Internet of Things leads to routing table explosion. An inexpensive approach for IP routing table lookup is required against ever growing size of the Internet. We contribute by a fast and scalable software routing lookup algorithm based on a multiway trie, called *Poptrie*. Named after our approach to traversing the tree, it leverages the population count instruction on bit-vector indices for the descendant nodes to compress the data structure within the CPU cache. Poptrie outperforms the state-of-the-art technologies, Tree BitMap, DXR and SAIL, in *all* of the evaluations using random and real destination queries on 35 routing tables, including the real global tier-1 ISP's full-route routing table. Poptrie peaks between 174 and over 240 Million lookups per second (Mlps) with a single core and tables with 500–800k routes, consistently 4–578% faster than all competing algorithms in all the tests we ran. We provide the comprehensive performance evaluation, remarkably with the CPU cycle analysis. This paper shows the suitability of Poptrie in the future Internet including IPv6, where a larger route table is expected with longer prefixes.

CCS Concepts

•**Networks** → **Routers;** *Network algorithms;*

Keywords

IP routing table lookup; longest prefix match; trie

SIGCOMM '15, August 17-21, 2015, London, United Kingdom

ACM ISBN 978-1-4503-3542-3/15/08.

DOI: http://dx.doi.org/10.1145/2785956.2787474

1. INTRODUCTION

The fundamental functionality to support high speed communications on the Internet are increasing in importance with the ever-growing traffic size and due to our daily life being highly dependent on the Internet. One of the key technologies is IP routing table lookup: It needs to be extremely fast since the peak traffic size in an Internet core router is multiple hundreds of gigabits per second (Gbps). Ternary Content Addressable Memory (TCAM) performs high speed IP routing table lookup in the Internet core routers. However, departure from a TCAM is an approach worth considering for the two reasons: First, TCAM has issues in power consumption and heat. Second, the advent of Network Functions Virtualization (NFV) [6] may make the use of TCAM impossible, since the virtualized network functions are currently implemented in software without TCAMs. Therefore, it is desired to implement a software high-speed IP router only with general purpose computers; i.e., personal computers (PCs), or commercial off-the-shelf (COTS) devices.

For a long period of time, IP routing table lookup has been the bottleneck [11] in the performance of the software IP forwarding using COTS devices. It is a challenging problem [34, 31] because: 1) the size of the routing table is large and keeps growing (the number of BGP full routes is beyond 500K), 2) it requires a specific compute-intensive processing step called "longest prefix matching," and 3) high-speed communication links require high speed processing (148.8 million lookups per second (Mlps) for wire-rate IP packet forwarding on 100 Gigabit Ethernet (GbE) of minimum-size packets).

Recently, we see a significant improvement in the performance of software routers. There are two approaches; one is expecting the use of specific hardware such as graphics processing unit (GPU). Such hardware inherits, however, similar issues that TCAM has, such as heat and power consumption. The other is a pure software algorithm approach, where we assume the use of the commodity CPUs. A trend of this approach is reducing the memory footprint of the data structure of IP routing table to maximize the benefit of CPU cache [38,