Cloud Gateways Need Better Top-*n* Detection for Fine-grained Hardware Offloading

FUTUREWEI Technologies
UNIVERSITY OF OREGON | VIRGINIA TECH
Chris Misa, Longzhang Fu, Riff Jiang, Liguang Xie

Overview

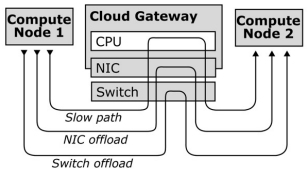

Figure 1: Example of a hybrid software/ hardware cloud gateway showing possible hardware offloading paths (SmartNIC and switch ASIC).

- Cloud data center networks rely on hybrid hardware/software "cloud gateways" [1] for network functions as shown in Fig. 1.
- Software must detect the top "heaviest" flows to offload to hardware.
  - Hardware resources limited to *n* flows.

The (Approximate) Top-n Problem

- Efficiently detecting top-n flows in software is highly challenging.
  - Cloud data center must process 10s of millions of flows, 100 Gbps traffic rates.
  - Critical to reduce CPU cycles, memory, and latency of detection.
    - Otherwise, benefits of offloading may be out-weighted by overheads.
- Need to solve "approximate top-n" problem (Fig. 2)

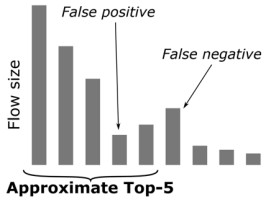

Figure 2: Simple example of the approximate top-*n* problem with *n* = 5.

Our observation: we can adapt approximate top-n solutions from network monitoring efforts to the cloud gateway offloading problem.

- High-level idea
  - Only maintain table of approximate top-n flows;
  - Strategically evict smaller flows such that the table approaches actual top-n.
- Based on "Space Saving" algorithm [2] shown in Fig. 3.

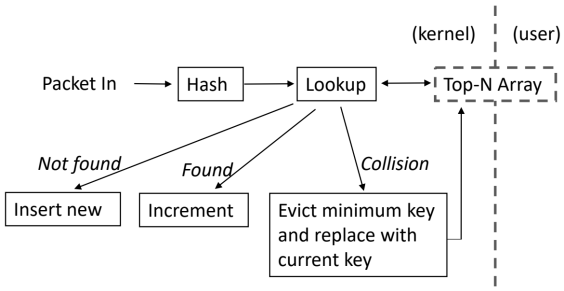

Figure 3: Update logic of the original "Space Saving" algorithm for detecting the top-*n* flows in O(*n*) space and constant per-packet time.

Citations

[1] Wang, Yanshu, et al. "Elixir: A High-performance and Low-cost Approach to Managing Hardware/Software Hybrid Flow Tables Considering Flow Burstiness." *USENIX NSDI, 2022.*
[2] Metwally, Ahmed, Divyakant Agrawal, and Amr El Abbadi. "Efficient computation of frequent and top-k elements in data streams." *International conference on database theory*, 2005.
[3] Yang, Tong, et al. "Elastic sketch: Adaptive and fast network-wide measurements." *SIGCOMM*, 2018.
[4] Sivaraman, Vibhaalakshmi, et al. "Heavy-hitter detection entirely in the data plane." *Symposium on SDN Research*, 2017.
[5] Yang, Tong, et al. "HeavyKeeper: An Accurate Algorithm for Finding Top-k Elephant Flows." *Transactions on Networking*, 2019.

Evaluation & Preliminary Results

- We consider a simple cloud gateway with eBPF/XDP software "slow-path" and SmartNIC hardware "fast-path" as shown in Fig. 4.
- We implement four state-of-the-art approx. top-n algorithms [3,4,5].
- Compare with packet-sampling baseline (e.g., proposed in Elixir [1]).

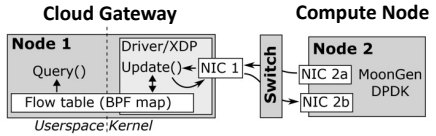

Figure 4: Testbed used to evaluate top-*n* implementations in simplified XDP-based cloud gateway.

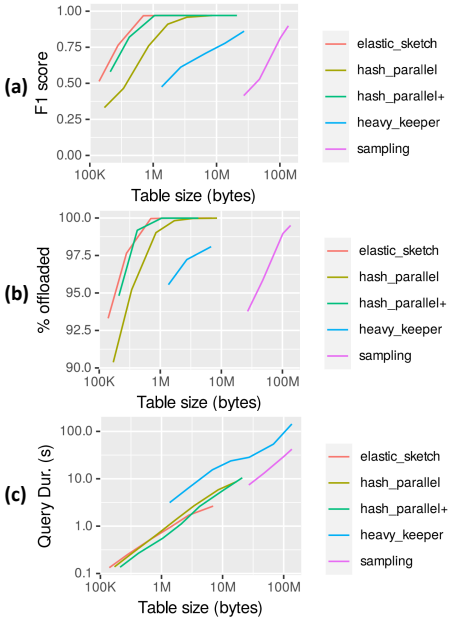

Figure 5: Impact of host memory (Table size) on top-n accuracy (a) volume of traffic offloaded compared to ground truth (b) and duration of querying top-n from XDP layer.

- Approx. top-n algorithms achieve near-perfect accuracy with ~2 orders of magnitude less memory compared to sampling (Fig. 5a).
- This translates into nearly as good of offloading performance as an exact top-n computation (Fig. 5b).
- Querying the top-n result from kernel takes around a second for adequately-sized tables (Fig 5c).
- CPU utilization remains low (~25%) even up to 30 Mpps (Fig 6).

Figure 6: Impact of forwarding rate on CPU utilization.

Future Work

- Monitor both software and hardware layers for longer-term maintenance.
- Deal with flow burstiness and dynamics.
- Extend beyond top-*n* problem to other hybrid hardware/soft traffic monitoring tasks.