实验方案

赵宗义

2019 年 11 月 13 日

Exp #84491

- Use the simulator of HashFlow implemented in python.

- Set the memory size to be 1 MB, so it can accommodate around 55K flow records.

- Select 10 trace files from each of the four traces.

- Initiate 50K flows from each trace file.

- Increase the depth of HashFlow from 1 to 4.

- Count the packets processed by the simulator, i.e., the number of original packets plus the resubmitted packets.

Exp #84492

- Select a file from the CAIDA trace (equinix-nyc.dirA.20180315-125910.UTC.anon.pcap), extract the first 2.5 million packets, classify the TCP/UDP packets into flows, and then calculate the average size as well as the maximum size of the flows.

- Select a file from the HGC trace (20080415000.pcap), extract the first 2.5 million packets, classify the TCP/UDP packets into flows, and then calculate the average size as well as the maximum size of the flows.

- Calculate the average as well as maximum size of a file from China Telecom trace (nfcapd.201601022000).

- Calculate the average as well as maximum size of a file from Tsinghua campus trace (20140206-6).

Exp #84493

- Set the memory size to be 1MB.

- Increase the number of flows from 10K to 100K, in the step size of 10K.

- Use a trace file from CAIDA and HGC respectively.

- Use four versions of HashFlow. In the versions the number of buckets in the ancillary table is $0.25\times$, $0.5\times$, $1.0\times$, and $2\times$ respectively of the number of buckets in main table.

- Calculate the average relative error for flow size estimation.

Exp #84494

- Randomly select a file from the traces of ChinaTelecom, HGC, Tsinghua and CAIDA respectively.

- Extract 5 million packets from each trace file.

- Record the number of distinct flows in each trace file. Note that all the packets with the same source IP address, destination IP address, source port, destination port, and protocol belong to the same flow. The flow ID is the five tuple (srcip, dstip, srcport, dstport, protocol).

- Map the flow ID of each flow to a digest using a given hash function.

- Record the number of distinct digest for each trace file.

- Compare the number of distinct flows and the number of distinct digests for each trace file.