Space-Code Bloom Filter for Efficient Per-Flow Traffic Measurement

Abhishek Kumar, Jun (Jim) Xu
Networking and Telecommunications Group
College of Computing
Georgia Institute of Technology
{akumar,jx}@cc.gatech.edu

Jia Wang, Oliver Spatschek
AT&T Labs - Research
{jiawang,spatsch}@research.att.com

Li (Erran) Li
Bell Laboratories
erranlli@bell-labs.com

Overview

Problem Statement

Per-Flow Counters and Other Approaches

Solution Architecture

Balls in bins - Counting probabilistically.

The Bloom Filter - An efficient implementation of bins.

A Multi-Resolution balls-in-bins counter

Evaluation

Problem Statement

Problem: To keep track of the total number of packets belonging to **each** flow at a high speed link.

Applications like traffic characterization, anomaly detection, etc., need to know the size of all flows.

Definition of Flow: All packets with the same flow-label.
The flow-label can be defined as any combination of fields from the IP header, e.g <Source IP, source Port, Dest. IP, Dest. Port, Protocol>.

Per Flow Counters - shortcomings.

- Majority of the packets belong to large flows, yet a majority of the flows are small.
- Large number of wide counters.
- High cost of storing flow-labels with corresponding counters.
- Amortization of worst case behavior is difficult.

Other Approaches

Sampling Sample packets with a fixed probability p and trace/process headers of sampled packets.

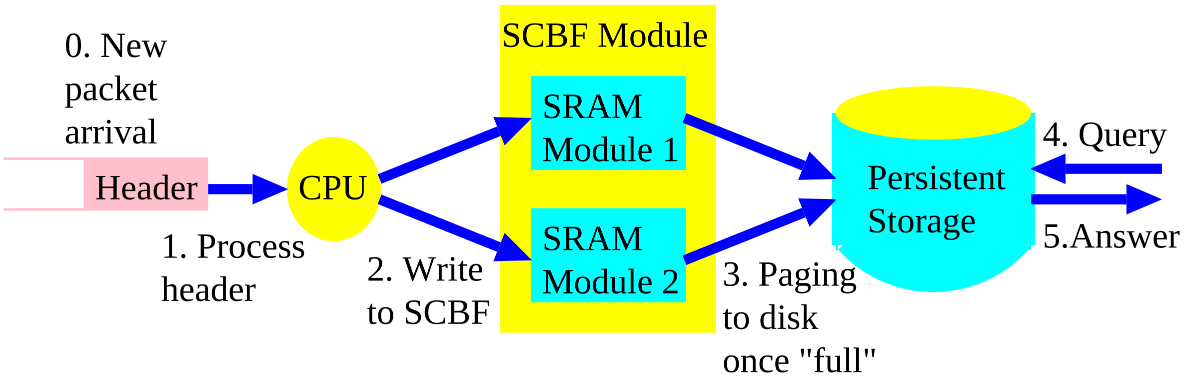
- Flow-sizes can be inferred from sampled data.
- Space-intensive.
- Inaccurate, especially for small flows.[Hohn and Veitch, IMC 2003]

Keep track of elephants

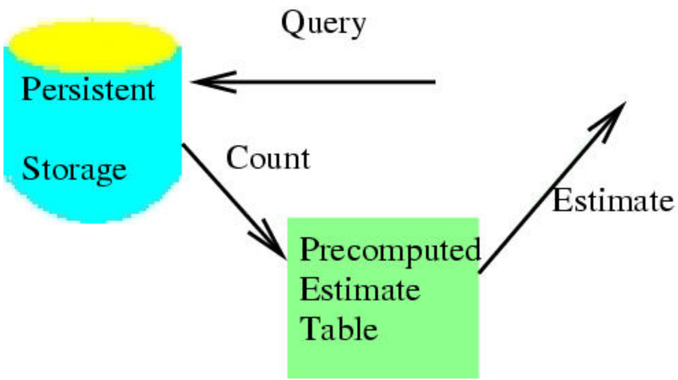
- Fast algorithm to filter packets from large flows. [Estan and Varghese, 2002]
- Maintain counters for large flows only.
- Success in tracking the largest few flows (e.g. carrying $\geq 1\%$ of the total traffic) with limited memory.

Architecture of Space-Code Bloom Filter – Insertion Module

- Measurement proceeds in epochs (e.g. 10 second).
- Maintain an aggregate synopsis data-structure.
- Update the data-structure on every packet arrival.
- Write-only data structure → fast updates, low hardware complexity.
- Copies of the synopsis are paged to disk periodically.



Architecture of Space-Code Bloom Filter (SCBF) – Estimation Module

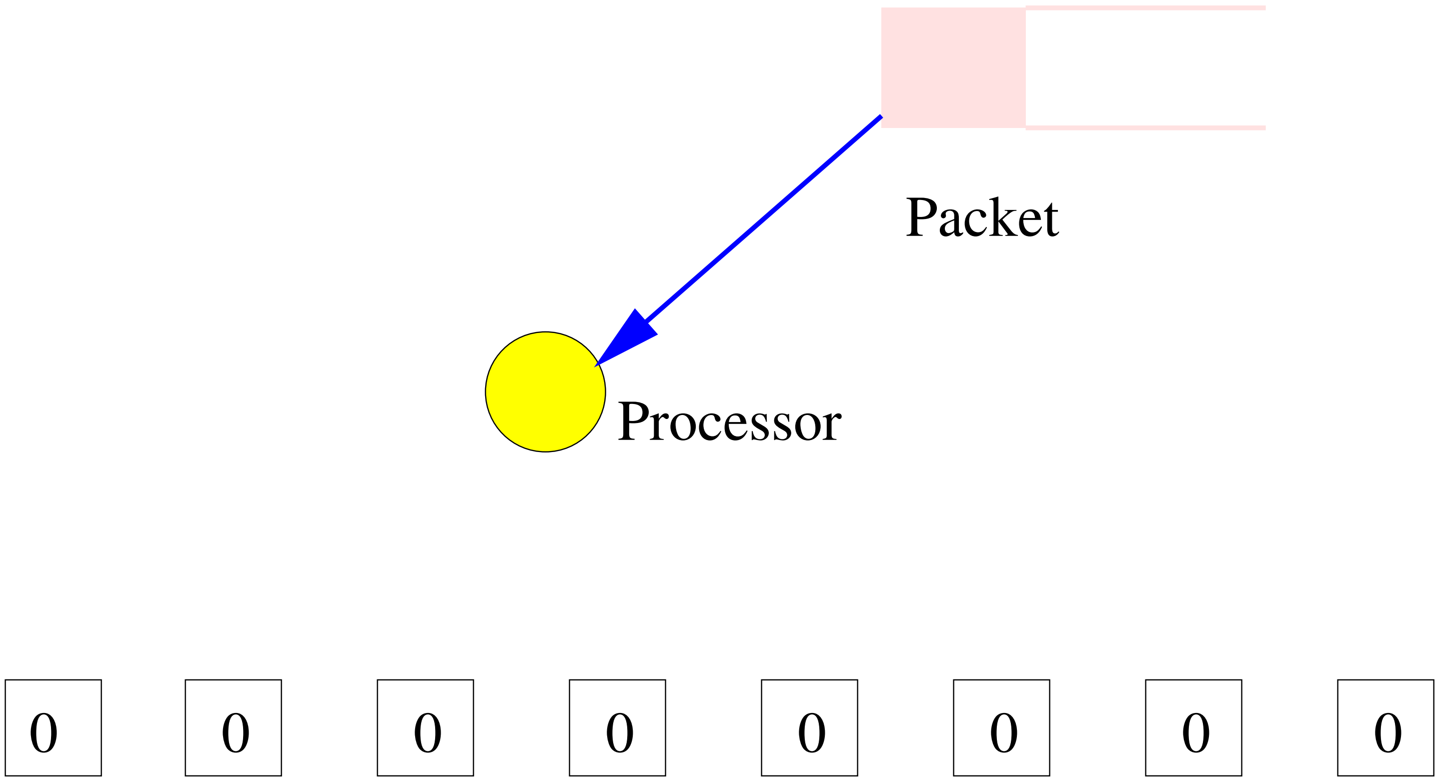


- Queries provide a flow-label and ask for its size.
- Obtain a "count" from the data-structure and then lookup a precomputed table to return approximate size of the flow.
- This provides approximate estimates that have low relative error with high probability.

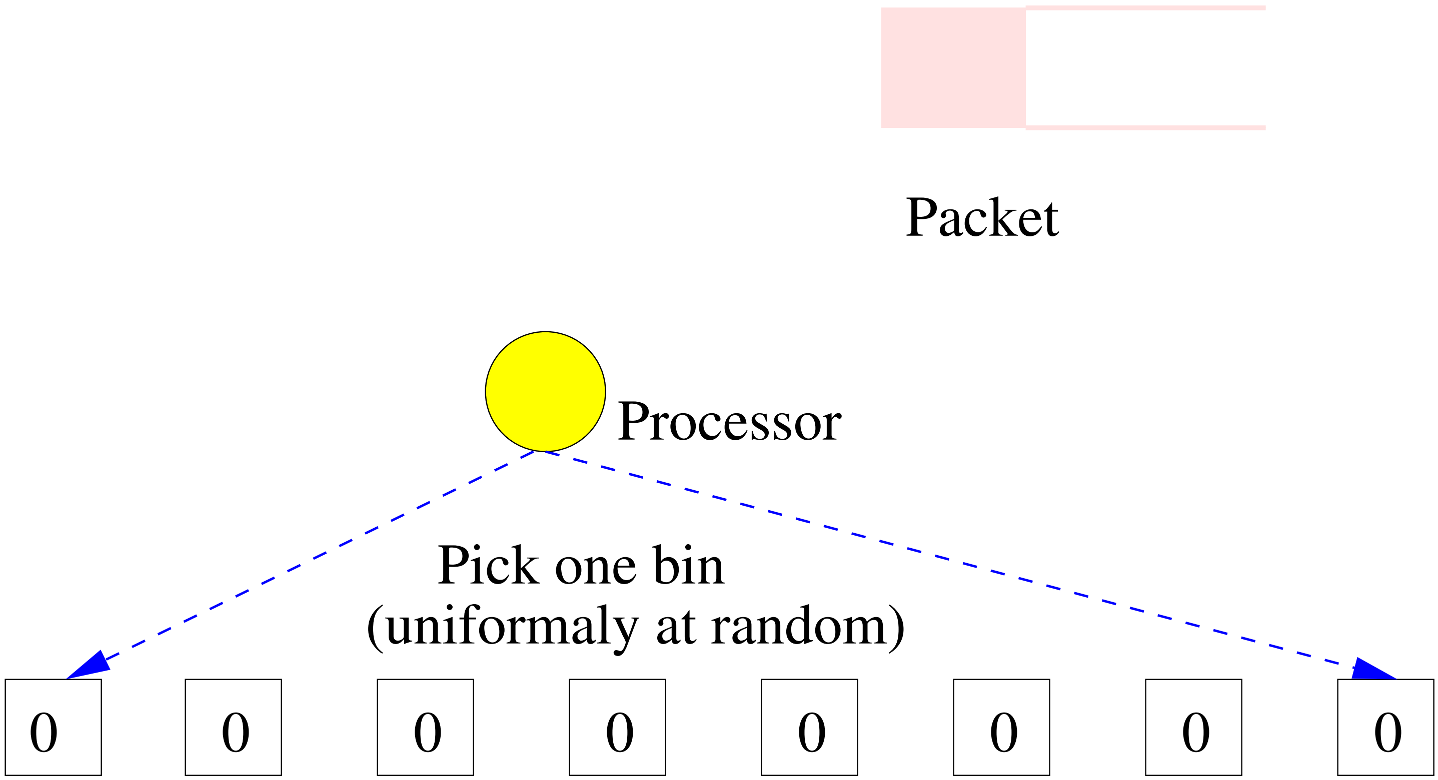
Balls in Bins

0	0	0	0	0	0	0	0

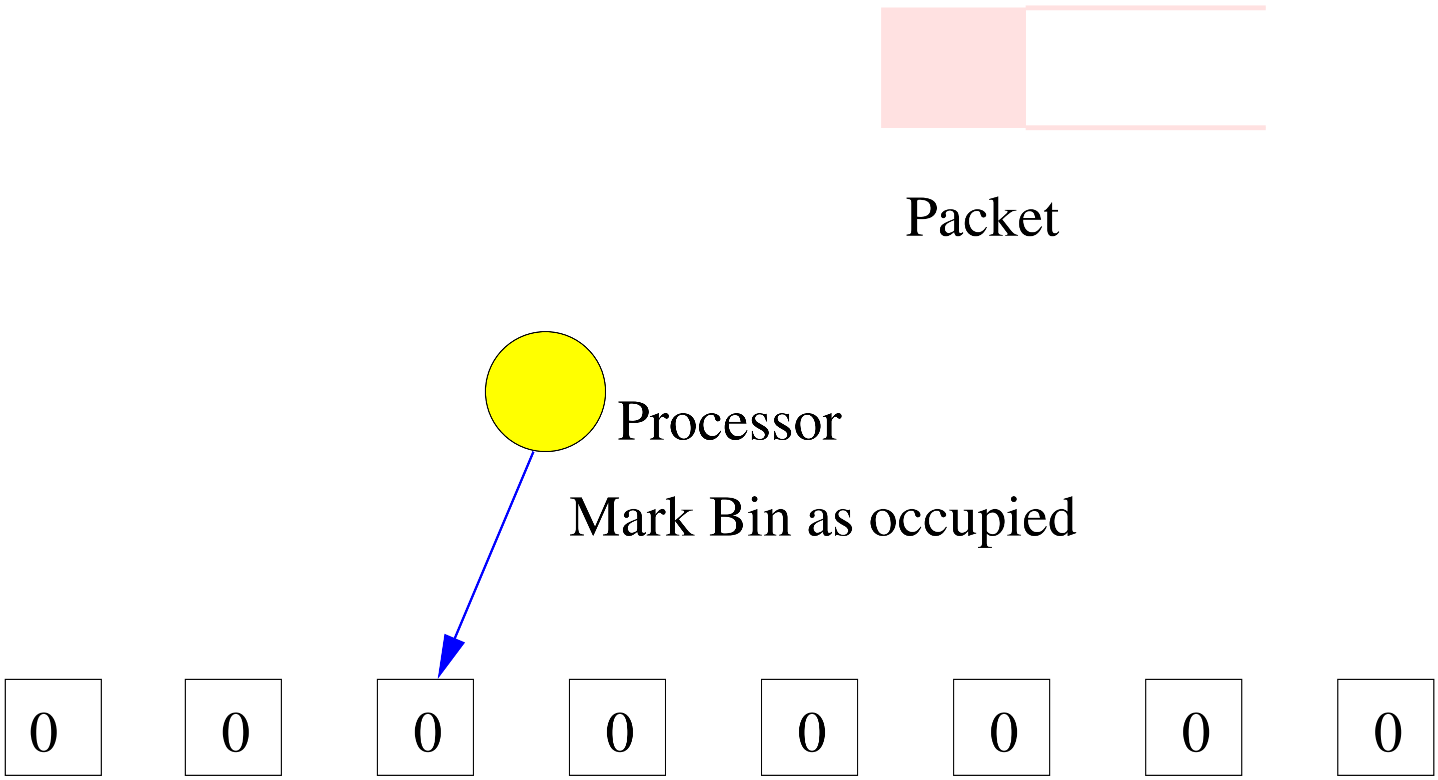
Balls in Bins



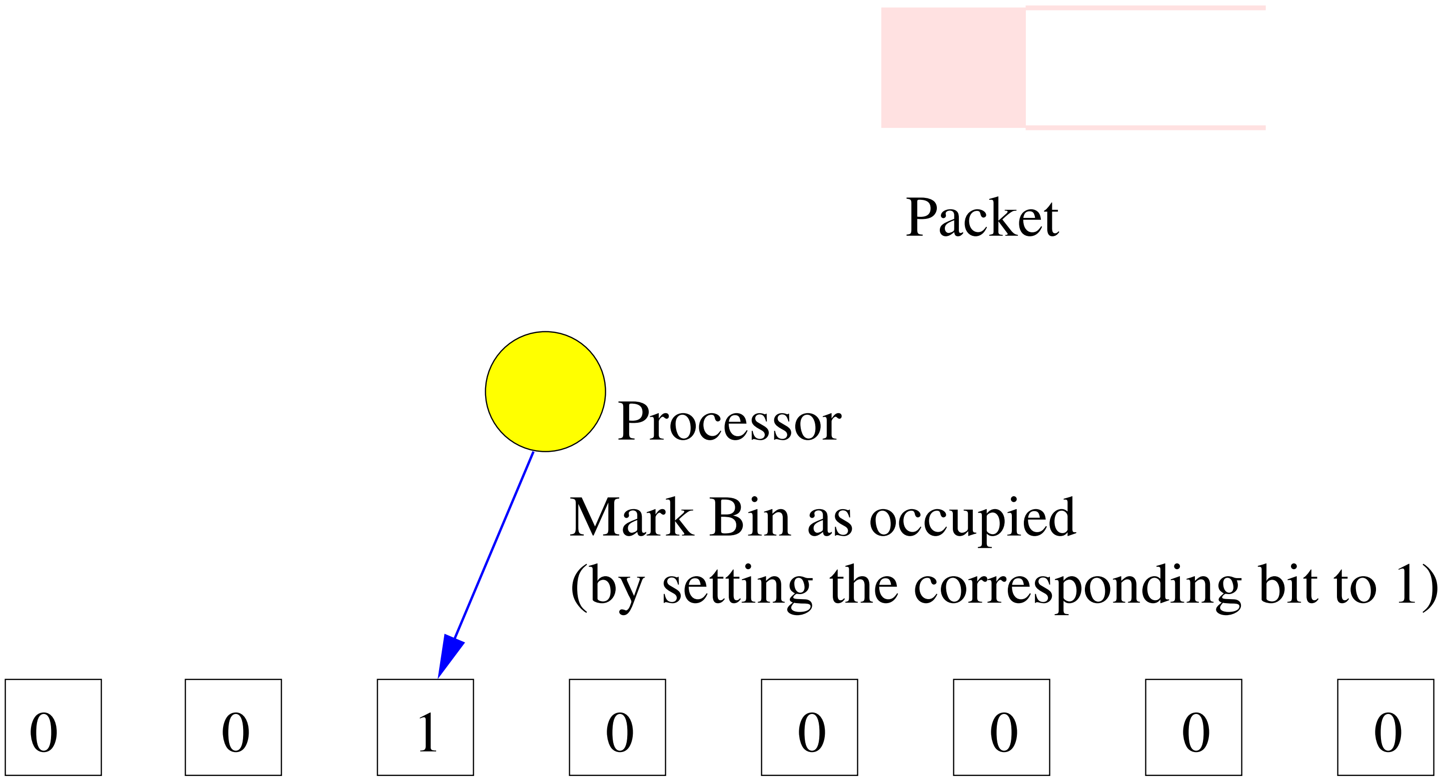
Balls in Bins



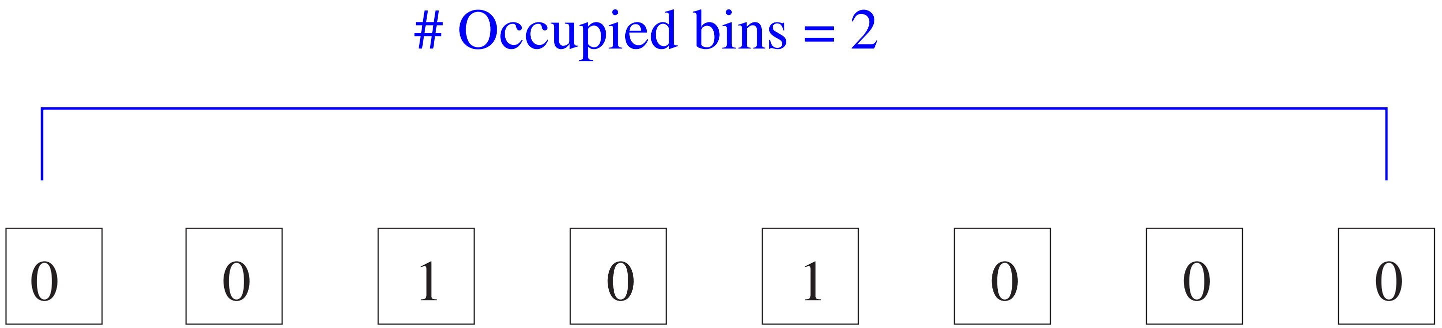
Balls in Bins



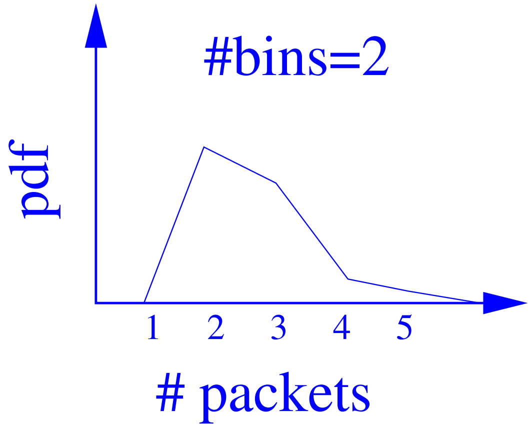
Balls in Bins

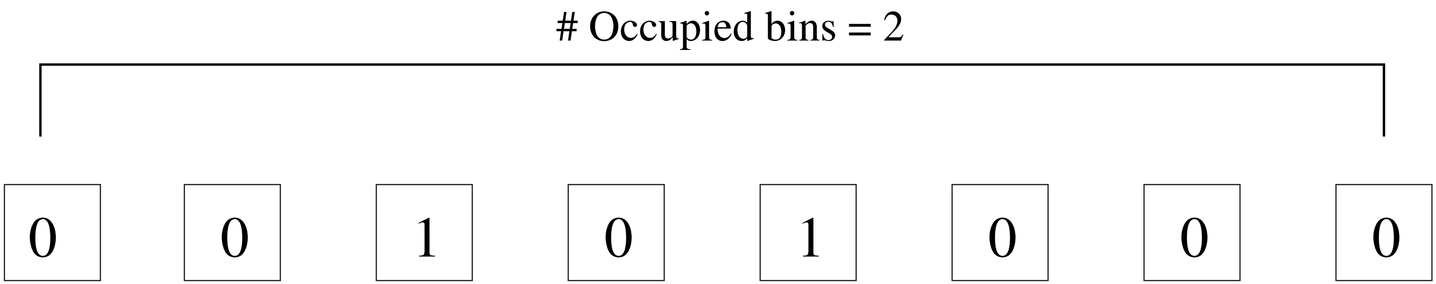


Balls in Bins - Estimation

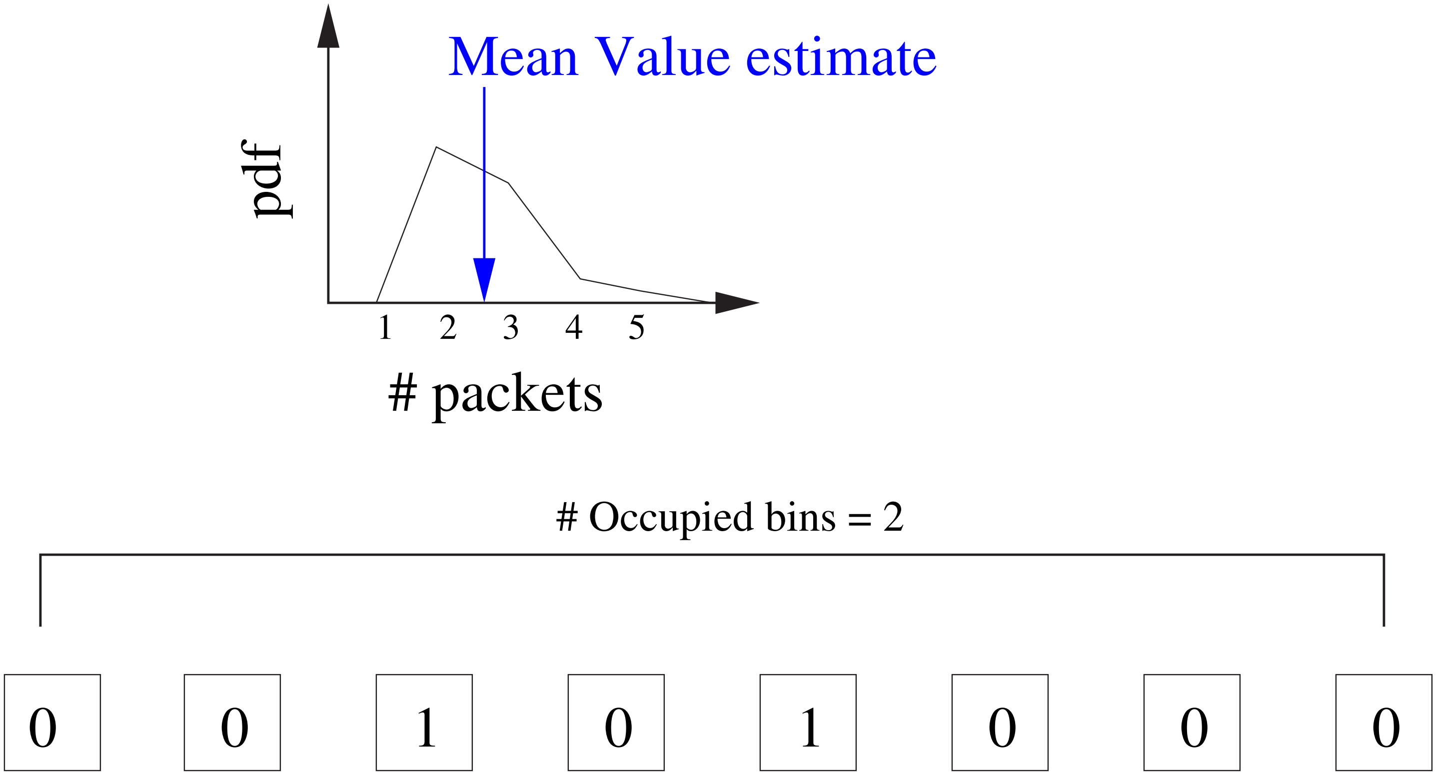


Balls in Bins - Estimation

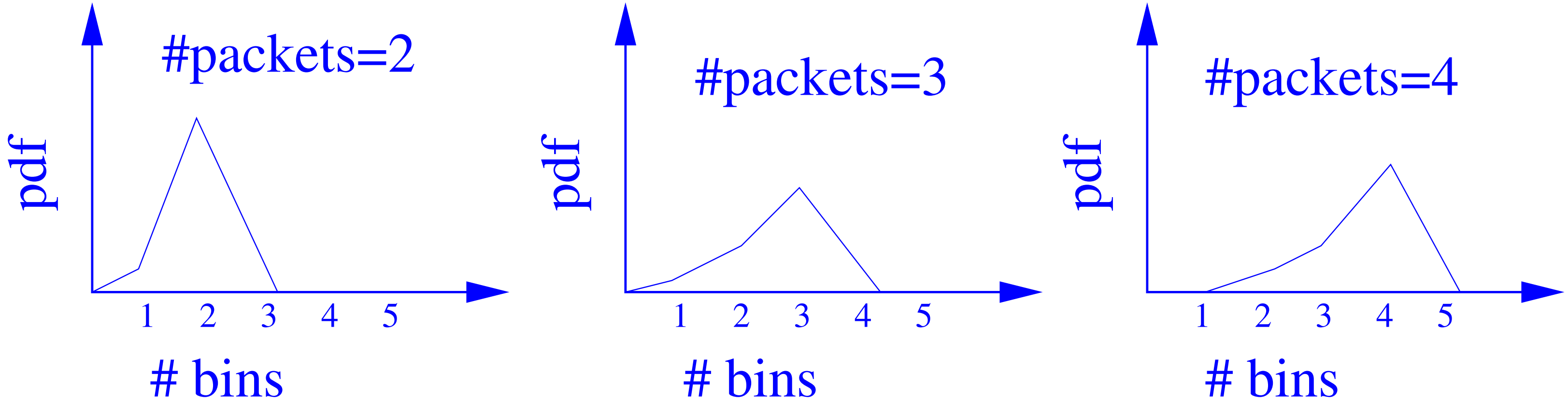


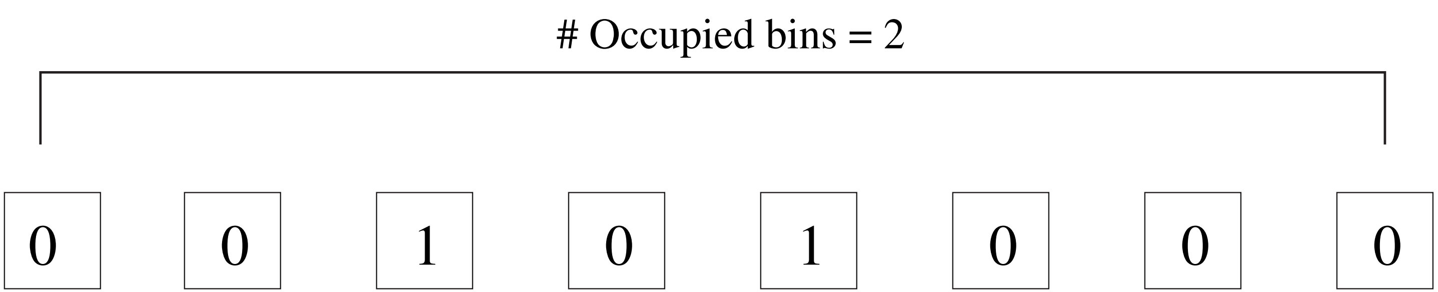


Balls in Bins - Estimation

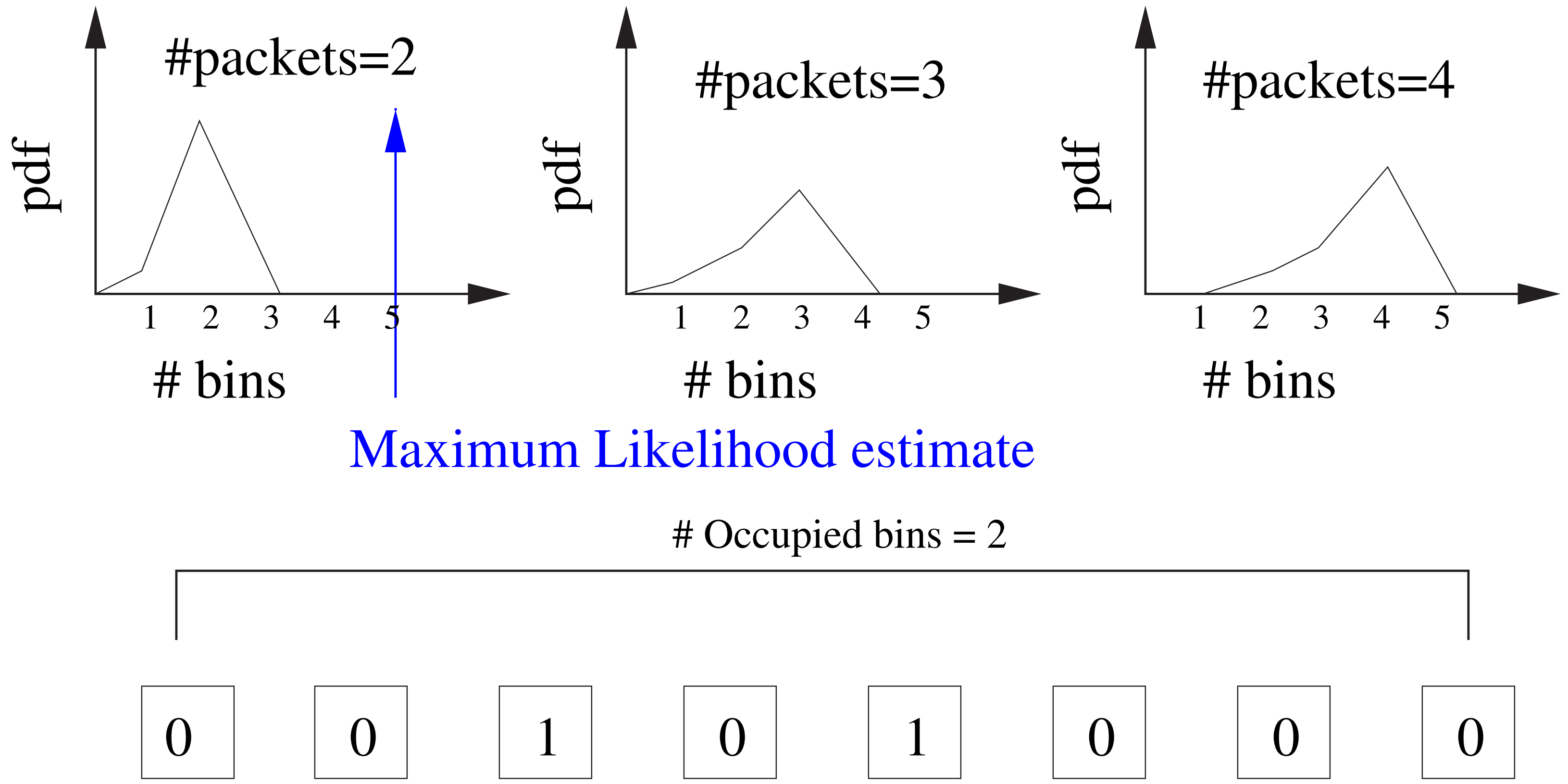


Balls in Bins - Estimation

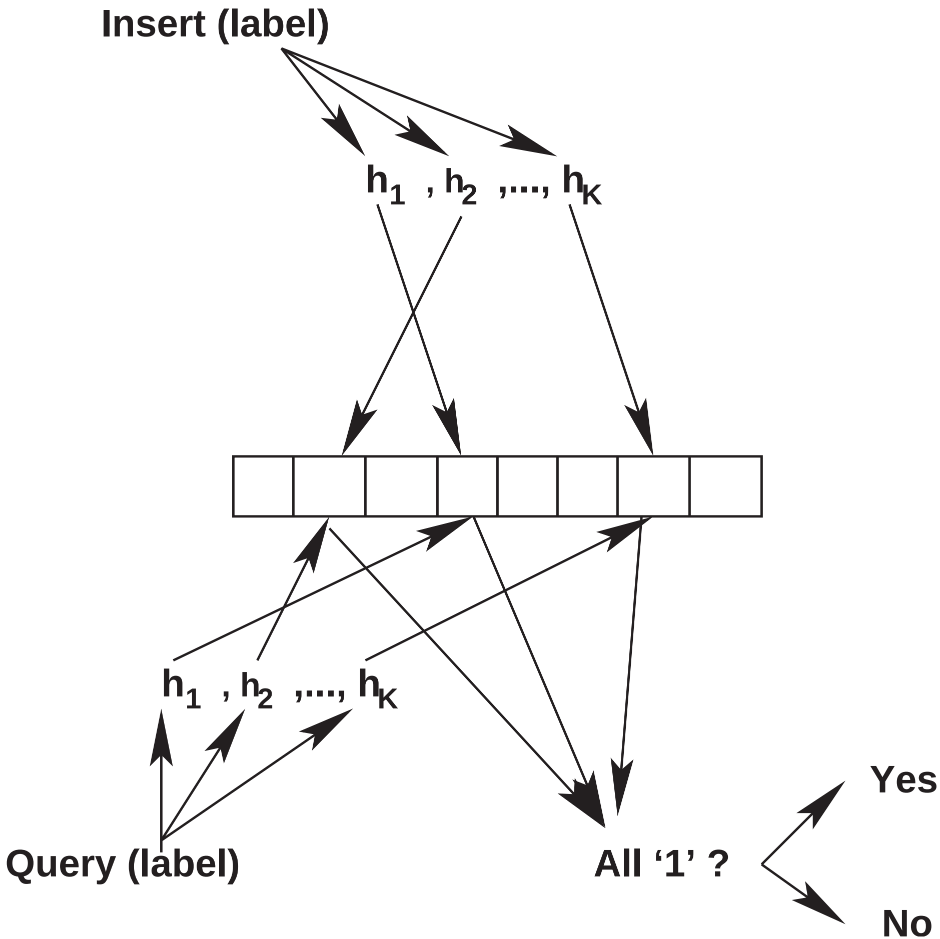




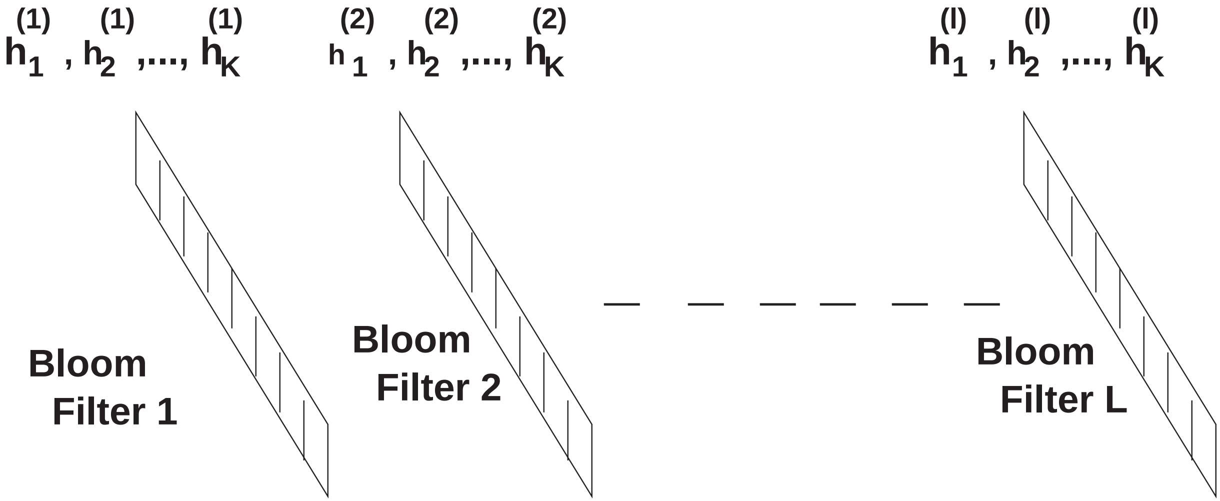
Balls in Bins - Estimation



Using Bloom Filters to Implement Bins

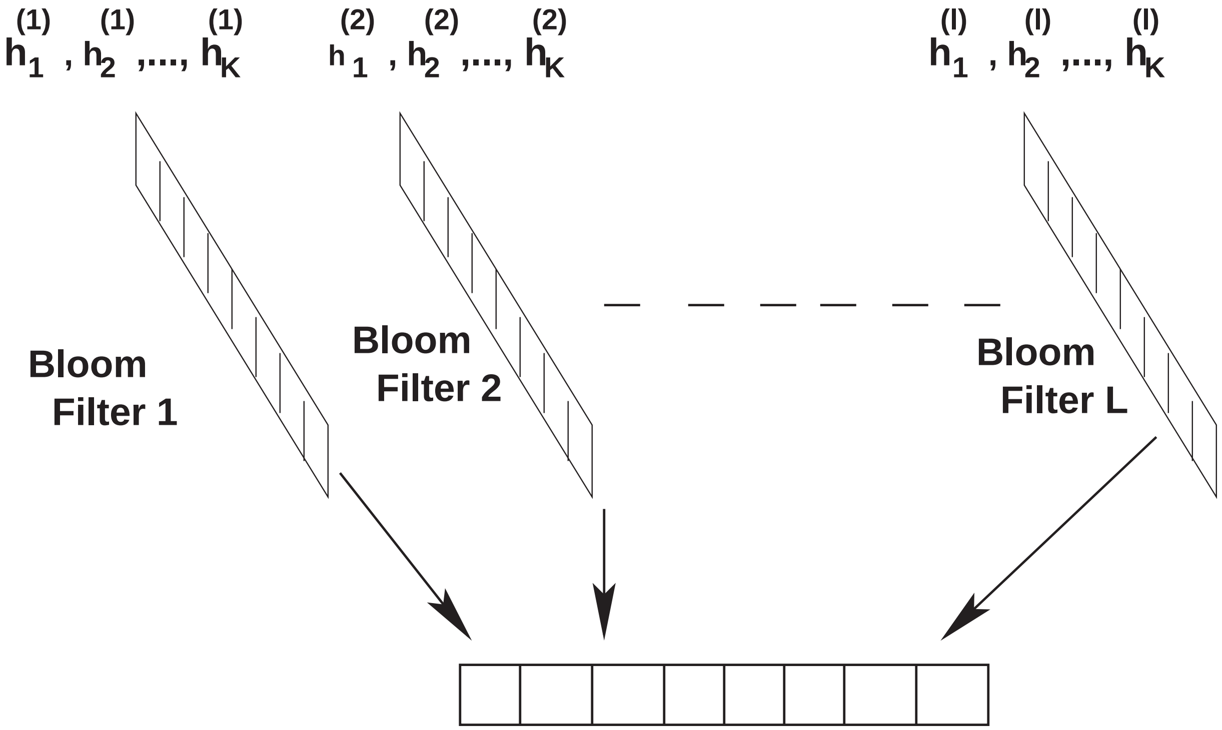


Use of multiple Bloom Filters to implement Multiple Bins

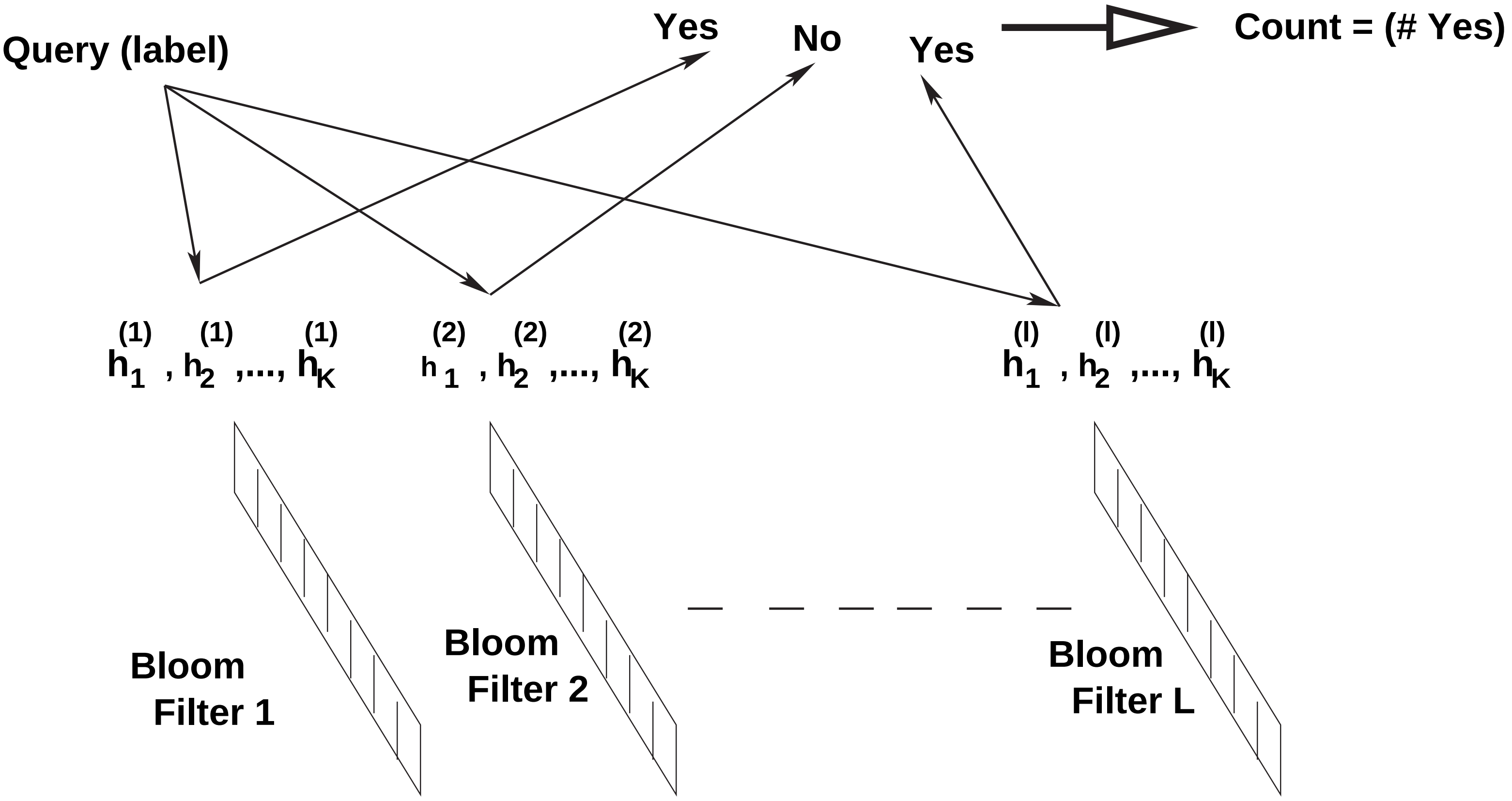


- Insert each ball (packet) in one of l bins (Bloom filters), chosen uniformly at random.
- While querying, count number of bins (Bloom filters) in which one or more balls (packets) were inserted and estimate the size of the flow from this count.

Space Code Bloom Filter - Multiple bins in the same Bloom Filter



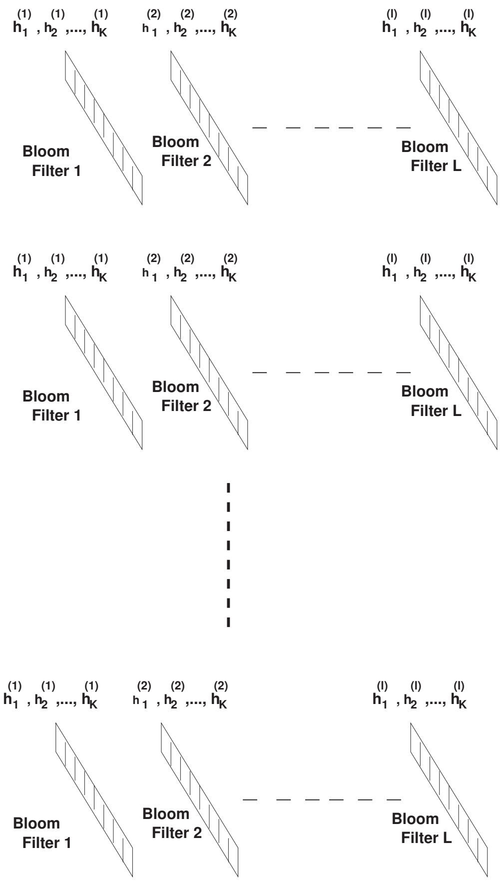
Querying an SCBF



Extending the range of SCBF

- An SCBF of size l fills up after about $(l \ln l)$ insertions (result from the classic coupon collector's problem).
- So we can't use it for estimating flow sizes much larger than $(l \ln l)$.
- Objective – To provide a constant relative error, irrespective of flow-size.
- Solution – Use multiple SCBFs operating at different resolutions.
- Small flows will be captured by SCBFs of finer resolution.
- Large flows will be captured by SCBFs of coarse resolutions.

Multi-Resolution SCBF



SCBF 1 –– Sampling rate 1

SCBF 2 –– Sampling rate 1/4

SCBF r –– Sampling rate $1/4^{r-1}$

Extending the estimation mechanisms to the Multi-resolution SCBF

Let the set of observations from the r filters be $\Theta = \theta_1, \theta_2, \cdots, \theta_r$.

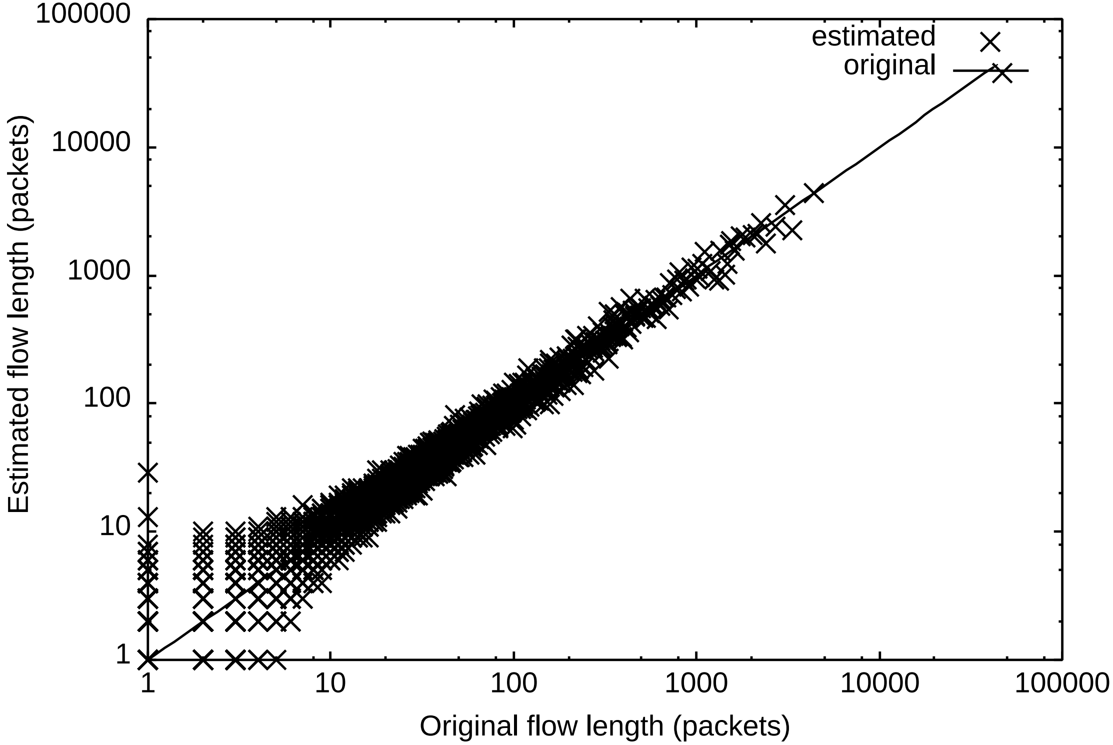
Maximum Likelihood Estimation

- Replace the MLE by a joint MLE over the set of r observations.
- In practice, choosing three "most relevant" filters is enough.

Mean Value Estimation

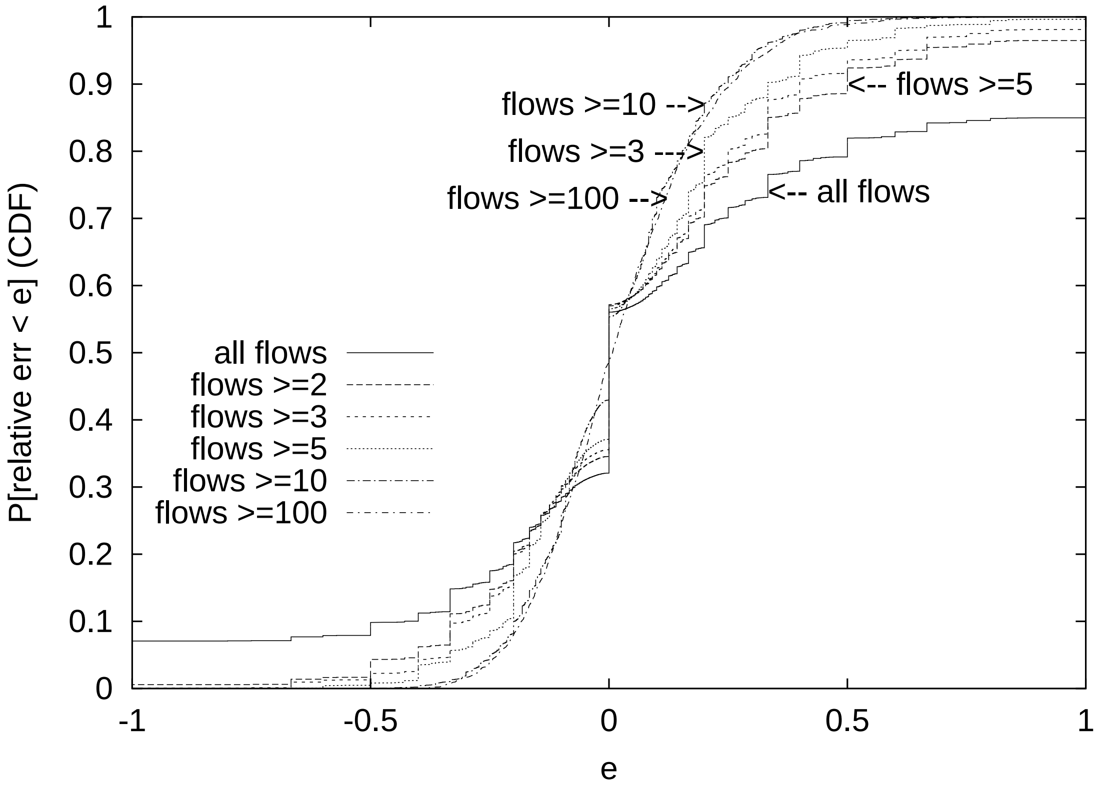
- Choose the most relevant filter and scale up its estimate by the inverse of the sampling rate.

Accuracy of SCBF



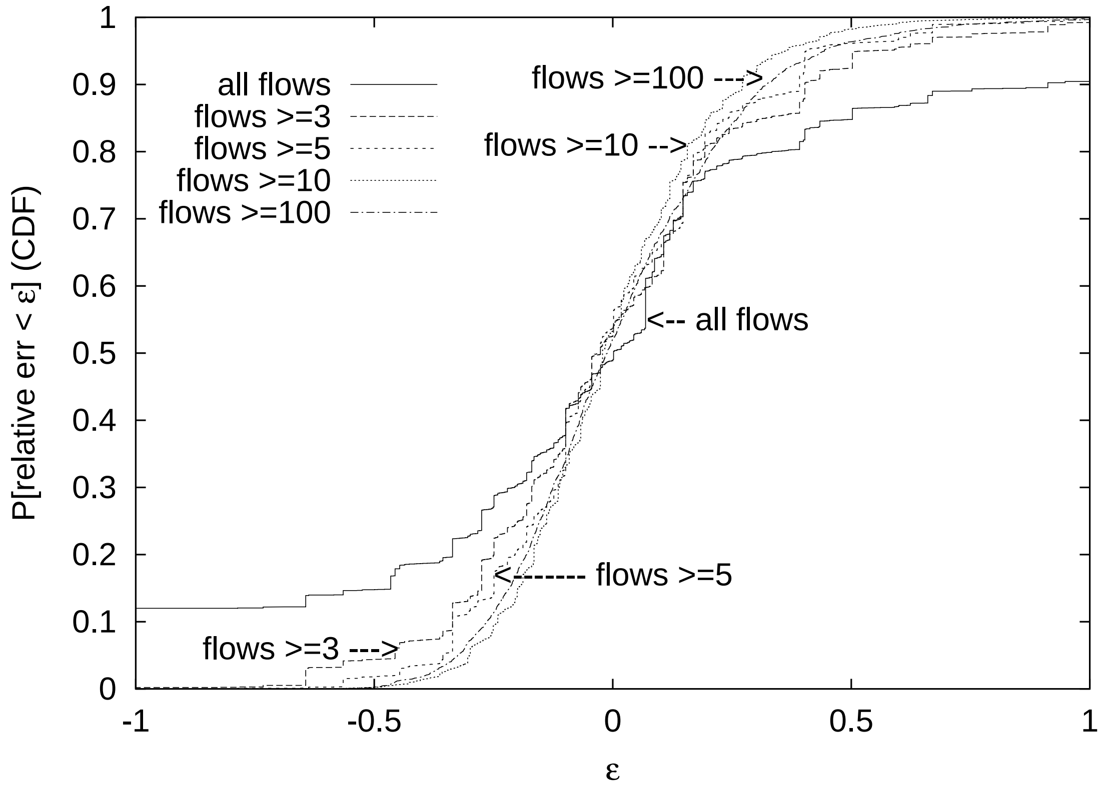
Original vs. estimated flow size. Note that both axes are on logscale.

Accuracy of SCBF using Maximum Likelihood Estimation



CDF of relative error for flows of various size

Accuracy of SCBF using Mean Value Estimation



The cumulative distribution of relative error in flow size estimation using MVE with 32 groups.

Performance of SCBF - complexities

- Computational complexity – compute 5 hash-functions and write 5 bits per packet.
- Space complexity – 4 bits of storage required for each packet.
- Can operate at OC768 (40 Gbps) with 5 ns SRAM.
- More than 80% responses are within $\pm 25\%$ of the actual value.

Conclusions

- Space-Code Bloom Filters can track the **approximate** size of every flow.
- Per-flow accounting without per-flow state.
- The relative error in approximation is same for all flow-sizes.
- Very fast (up to OC768) implementations possible due to “write-only” nature of updates.
- Design parameters of SCBF can be tuned to trade storage space and CPU cycles for accuracy.

Questions ???

Identifying the most relevant filter for estimation

- For an SCBF with l groups, θ of which are matched by an item x, it would take about $\frac{l}{l-\theta}$ insertions on the average to match another unmatched group, and increase the observation to $\theta+1$.
- We know from the coupon collector's problem that the total number of insertions required to cause the observation θ is $(\frac{l}{l} + \frac{l}{l-1} + \cdots + \frac{l}{l-\theta+1})$.
- Thus the *relative incremental inaccuracy* of this observation is $\frac{l}{(l-\theta)}/(\frac{l}{l} + \frac{l}{l-1} + \cdots + \frac{l}{l-\theta+1})$.

Identifying the most relevant filter for estimation