

上报带宽与数据结构参数设计说明文档

1. 数据平面上报至控制平面的带宽测算

在模拟实验中，数据平面hashtable平均上报约 267256.5 条记录，blooms和cmsketch平均上报 47752.4条记录。即使在最理想情况下（没有误报），也需要上报约 120,000 条数据，因此实际上仅多上报了约 200,000 条。

每条上报记录包含以下字段：

- 流 ID (Flow ID)：13 字节（例如使用五元组编码）；
- 端口 ID (Port ID)：2 字节；
- 上报层级标识：2 字节。
其中标识部分：hashtable为存储的流大小，最高4位为0000；bloom与cmsketch以位图形式存储，如最高4位为0001，标识为cmsketch上报，0010标识为第三个bloom上报。

如果有最小数据包大小限制，如每条记录至少需要 64 字节，则平均总上报数据量为：

$$(267256.5+47752.4) \times 64 \text{ B} = 20160569.6 \text{ B} \approx 20.2 \text{ MB}$$

结论：该上报数据总量仅为单次测量期间网络整体流量的极小部分，占带宽极低，可视为控制面通信成本非常小，不构成瓶颈。

2. 各层数据结构参数设置逻辑

2.1 Count-Min Sketch (CM) 配置逻辑

✓ 大流数量估算

- 根据数据集设定，大流的比例为 10%，对应约 12k 个流；
- 解码目标是尽可能还原这些大流的准确流量大小。

✓ 行数 d 的选择（哈希函数数量）

- 解码概率与哈希函数数量 d 强相关；
- 实验表明，当资源相同（d × w 固定）时，d = 3 相较 d = 2 能显著提高解码成功率；

- 继续增加 d (如 $d = 4$) 会显著增加资源开销, 不再划算;
- **最终选择: $d = 3$**

✓ 列数 w 的选择 (每行桶数)

- 实验发现, 当 $d = 3$ 时, 当 $w \geq 0.5 \times \text{流数 } n$, 可以以较高概率解码出大部分流;
- 取 $n \approx 12000$, $0.5n \approx 6000$;
- 通过实验, 综合内存开销与准确率, 选择 **$w = 10000$** 为平衡点;

✓ bit-width 的选择

- 每个桶的 bit-width 直接影响空间占用;
- 通过大流包数分析, 625万的包经过最少64个port拆分, 再因为hashtable中阈值设定 $2^8=256$ 进行压缩, 所以进入cm的单流大小不超过2g, 但考虑到充分的冲突预留 (多个流进入一个counter), 所以每个桶的bit-width设为12bits。
- **空间占用:**

$$w \times d \times \text{bit-width} = 10000 \times 3 \times 12 = 360,000 \text{ bits} \approx 45 \text{ KB}$$

2.2 多级 Bloom Filter 配置逻辑

✓ 设置多层结构的原因

- 单层 Bloom 过滤器即使设为340 bits, 也仅能识别 90% 的流键 (120k子流的情况);
- 为降低漏报风险、过滤更多小流, 采用多层 Bloom 策略;
- 每一层识别失败的流继续进入下一层, 最后进入 CM 的流是所有 Bloom 未能过滤的流;
- 因此采用三级结构, 可有效拦截大小为 1-2 的小流, 避免它们误入 CM。

✓ 层数 l 的选择

- 小流 (经过压缩拆分后) 大多大小为 1 或 2;
- 要屏蔽所有小流, 则需对 size=1、size=2 的流都尝试识别;
- 为避免 false positive 累积漏报, 需多设一层作为“缓冲”;
- **最终层数设置: $l = 3$**

✓ 每层哈希函数 k 的选择

- 因总 Bloom 预算约为 $700 \text{ KB} - 360 \text{ KB} = 340 \text{ KB}$;
- 若每层使用过多哈希函数 k , 会导致误报率下降但冲突上升;
- 为防止 Bloom 退化为哈希表, 同时保持计算简单, 选择 **$k = 2$** 为各层统一值。

✔ 每层数组长度 m 的选择

确定了k之后，我们就能通过公式和实验估算出每层m与漏报的概率。进入每层bloom的流n，可以近似为进入第i层bloom的流就是大小size大于等于i的流以及前一层第i-1层bloom漏报的流（假阳性）。

- 每层流量 n 逐层减少，因此我们近似各层 Bloom 空间 m 可逐层减半；
- 第一层需尽可能覆盖大部分流量，因此设置最大；
- 根据多次实验和流大小分布，多次实验观察每层上报流的数量，并且尽可能保证最后尽可能所有流键都能上报。

多轮实验后，最终选择：

m1 = 160,000 bits
m2 = 60,000 bits
m3 = 40,000 bits

3. 总体资源汇总

层级	类型	长度参数	哈希数	空间公式	空间合计
L0	HashTable	size=2000	1	size × 40 bits	2,000 × 40 bits = 80,000 bits (10 KB)
L1	Bloom	m1=160000	k1=2	m1 bits	160,000 bits (20 KB)
L2	Bloom	m2=60000	k2=2	m2 bits	60,000 bits (7.5 KB)
L3	Bloom	m3=40000	k3=2	m3 bits	40,000 bits (5 KB)
L4	CM	d=3,w=10000,b=12	d=3	d×w×b bits	3×10,000×12 bits = 360,000 bits (45 KB)

总内存使用 ≈ 87.5 KB=700Kb