

HyperLogLog(HLL)

SILICON VALLEY ENGINEER

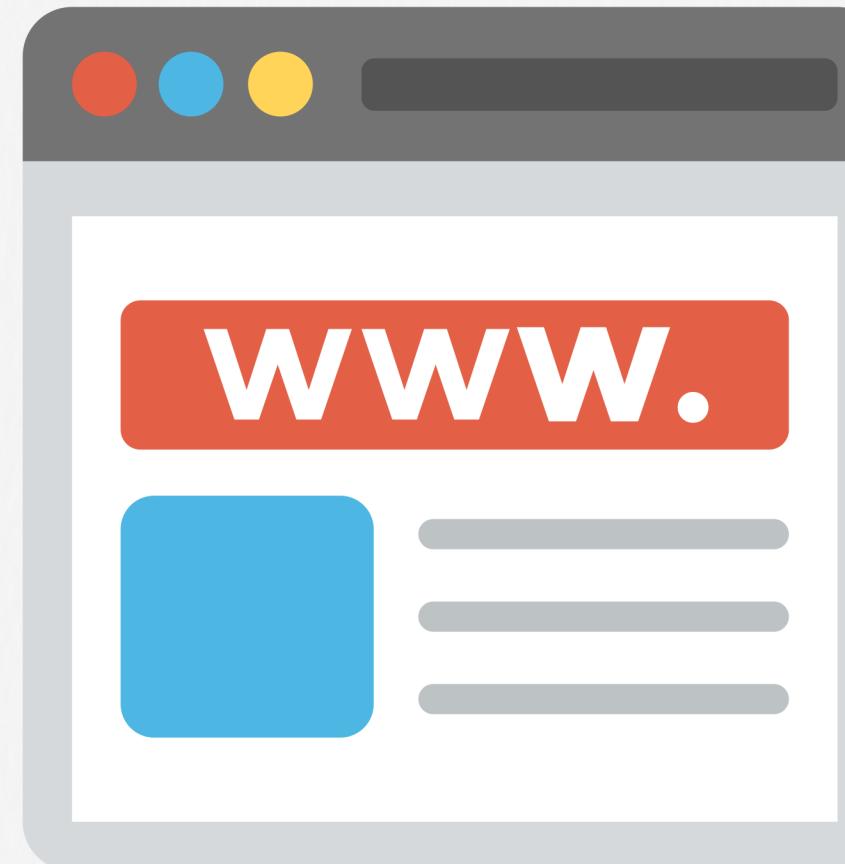
- Probabilistic data structure used in Redis for **approximating the cardinality** (count of unique elements) of a set
- **Fixed-Low Memoery footprint(12KB)** regardless of the number of elements in the set
- **Trade-off between memory usage and accuracy**, making it suitable for applications where **precise cardinality is not required but memory efficiency** is essential(Error rate is 0.81% - 992 ~ 1,008 out of 1,000)
- Use case
 - unique visitors to a website
 - unique ip address
 - unique items in a shopping cart
 - unique count with good security



log(log(list))

HyperLogLog(HLL) - Use case

SILICON VALLEY ENGINEER



Previous visitor set

Bob Alice

Bob

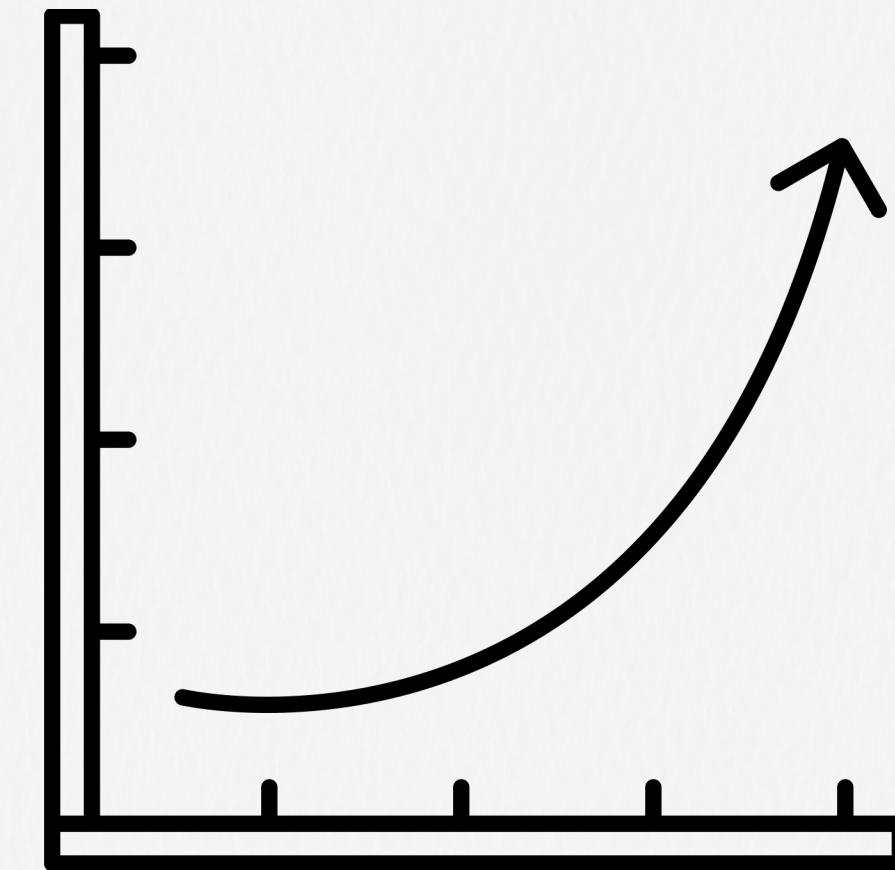


HyperLogLog(HLL) - Benefit

SILICON VALLEY ENGINEER

Previous visitor set

Bob Alice



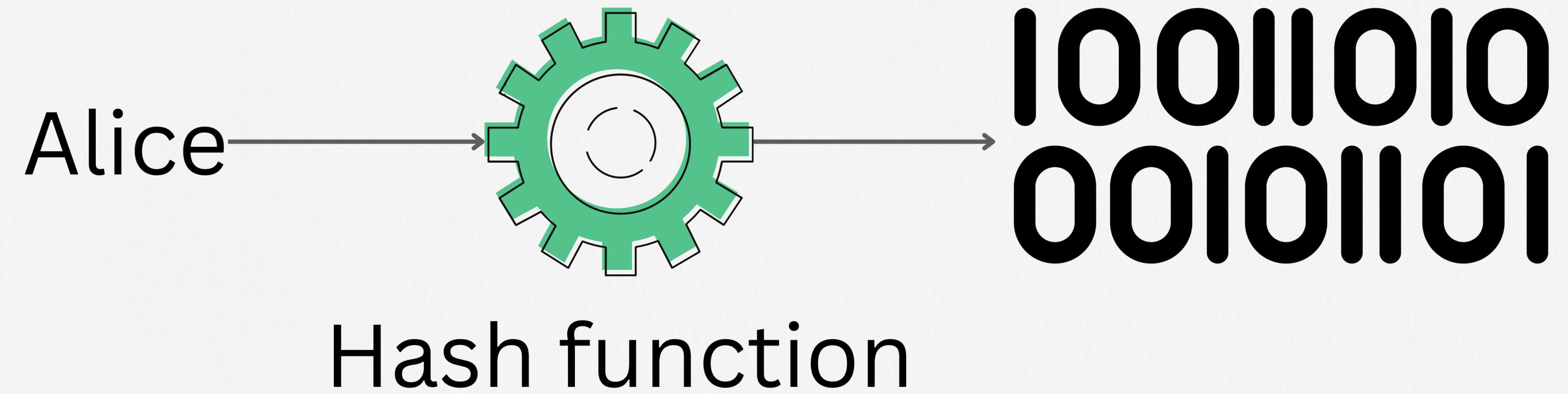
16bytes per username * 1000 = 16KB

16bytes per username * 10^6 = 16MB

16bytes per username * 10^9 = 16GB

HyperLogLog(HLL) - How it works

SILICON VALLEY ENGINEER



HyperLogLog(HLL) - Commands

SILICON VALLEY ENGINEER

- **PFADD**: add elements to the HyperLogLog
- **PFCOUNT**: estimate the approximate number of unique visitors in the HyperLogLog
- **PFMERGE**: merge multiple HyperLogLogs into a single HyperLogLog

