Summary - An Efficient PGP Keyserver without Prior Context

Bowen Song[1]
August 28, 2018

***Abstract*— This report is a part of an independent study for set and string reconciliation problems in distributed systems. The paper [1] is an implementation of set reconciliation protocol for PGP key servers using Inevitable Bloom Filter and Strata estimator.**

I. Introduction

This implementation considers the problem of synchronizing PGP key servers using the minimum amount of communication. The PGP key servers are in a distributed system without central authority. Each server contains a list of PGP keys generated by client requests. The key servers are looking for reconciling their list of keys periodically with the minimum amount of bandwidth utilization. The implementation is based on IBLT [2] as set reconciliation protocol and estimates symmetrical differences between key servers using Strata estimator.

II. Algorithm Overview

The protocol uses Strata estimator to assess the symmetrical differences between two key servers and reconcile their difference using one IBLT 3 times the size of Strata's estimation. Since the size of IBLT is directly related to the number of symmetrical differences it can recover, the implementation uses 3 as a constant to achieve high success rate reconciling the key servers.

The Strata estimator uses IBLT hierarchy to assess the symmetrical differences between two reconciling datasets. For each level of IBLT, the estimator samples dataset at a progressively increasing rate and insert the samples into an IBLT. The estimator gets decoded same way as an IBLT to estimate the number of symmetrical differences. Each level of Strata estimator samples the original set with probability 2^{-n}. As the n increases, the amount of samples decreases.

The implementation uses a Strata estimator with 2^n levels; Each level has a fixed sized IBLT containing an increasingly large number of samples. Each IBLT of Strata estimator uses 4 hash functions to place value in 4 different places, and the IBLT has a fixed size of 40 cells.

III. Algorithm Performance Analysis

The performance for the implementation evaluates PGP key servers with 20k keys and 200k keys and change the number of key differences from 1 to 1000. The amount of keys in a key server is not contributing any notable performance difference for either communication or computation analysis. This observation shows that the implementation is scalable without considering the size of the underlying database. For communication analysis, the IBLT has a linear increase in size for the number of bytes transmitted as the number of set symmetric differences increases, whereas the Strata estimator has a logarithmic increase. However, the performance data for Strata estimator is not enough to be conclusive.

IV. Conclusion

The implementation of set reconciliation protocol for PGP key servers is benchmark demonstration. It shows that the Strata estimator and IBLT can efficiently reconcile the difference between two large set of data, and the protocol scales linearly to the size of set symmetrical differences.

References

[1] R. Van Renesse, D. Dumitriu, V. Gough, and C. Thomas, "Efficient reconciliation and flow control for anti-entropy protocols," in *proceedings of the 2nd Workshop on Large-Scale Distributed Systems and Middleware*. ACM, 2008, p. 6.

[2] M. T. Goodrich and M. Mitzenmacher, "Invertible bloom lookup tables," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE, 2011, pp. 792–799.

[1] B. Song is with Department of Electrical and Computer Engineering, Boston University, Boston MA, `sbowen@bu.edu`