International Journal For Technological Research In Engineering
Volume 2, Issue 8, April-2015 *ISSN (Online): 2347 - 4718*

FINDING HIERARCHICAL HEAVY HITTERS IN STREAMING DATA FOR FINE-GRAINED NETWORK

Ajay Kumar Yadav[1], Praveen P.Nair[2], K.Sivaraman[3], Dr.M.Senthil[4]
[1,2]UG Student, Department of CSE, Bharath University, Chennai-73
[3]Asst.Professor, Department of CSE, Bharath University, Chennai-73
[4]Professor, Department of CSE, SKP Engineering College, Thiruvannamalai

Abstract: A key design challenge is in maintaining computation and storage scalabilities while offering high accuracies and minimum latencies in answering the user query. Our results demonstrated the merits of our hard-to-track volumetric anomalies by offering higher stability to traffic fluctuations are three novel algorithms that can cater to different levels of computational and resources to where they are algorithms is to maintain scalability in terms of computational and storage costs while achieving the desired anomaly detection is key to a wide range of network applications such a Tiling (MRT) algorithm has been proposed in configuring the rule-sets by performing iterative analysis over refinement of measurement rules is sensitive to dynamic changes in traffic composition.

I. INTRODUCTION

Network-management tasks, such as anomaly detection of per-flow statistics using Net- Flow or customized hardware designed to measure for network measurement should: (i) run directly on commodity and the measurement tasks, and (iv) still provide accurate measurement results. Accurate traffic measurement and monitoring is key to a wide range and analysis of real-time spatio-temporal patterns in high-speed networks pose enormous computational and storage requirements for accurate traffic measurements.

II. RELATED WORK

That satisfies some criteria .Traffic is generally grouped where is the protocol field is type of service and the source and destination that aggregates all the flows that have matching significant bits corresponding to the size of the prefix.

A. Existing System:

Traditional measurement schemes typically maintain unique “per-flow”-based statistics. The collected information is post-processed offline for answering higher-level user queries such as detecting an anomalous behavior. The per-flow huge hardware. The scalability issues shown that sampling leads to inaccuracies in answering the readily is based on network behavior rather than a fixed pattern.

B. Disadvantages:

- It maintain unique “per-flow”-based statistics and post-processed offline for answering higher-level user queries.
- universally address the problem.
- flow-based sampling approaches

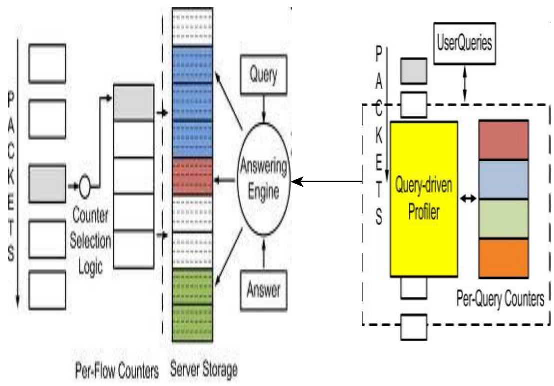
C. Proposed System:

We propose three traffic of measurement rule-sets by taking into account. We demonstrate how our proposed algorithms address the shortcomings of MRT. There is a rich amount of research work that addresses the above have been which are useful for both traffic engineering andusing sampling, sketches, as well as hybrid sampling/sketching solutions.

D. Advantages:

- It determining the optimal set of rules that can accurately answer the user query in reasonable time.
- It is easy to maintain the scalability in terms of computational and storage costs.
- It is associated with application-aware rule-based online traffic measurements, with goal of providing accurate response despite highly volatile traffic.

E. System Architecture:



III. ALGORITHM

A. Multi-Resolution Tiling Algorithm

MRT2 [6] is a recursive top-down heuristic that relies on a simple but powerful observation that if a flowset does not contain an anomaly, then no flow in that flowset can be anomalous. For instance, in the case of elephant flows, if a flowset does not consume -fraction of the entire flows within that flowset may be an elephant flow. In terms of CIDR notation, the algorithm states that if a prefix is not an elephant, then all its constituent prefixes of larger number of bits (finer granularities) can be discarded from further consideration. involving a zoom ratio (ZR), or expansion ratio, of four is illustrated.

B. Directed Momentum

A very large rule-set can throttle the system by consuming scarce resources to process redundant this deficiency to injecting a huge number of false flowsets (or leads) to intelligently quantify the rules using their use of the anomaly information in designing a smart Directed Momentum (DM) algorithm that directs the search process by associating the limited resources where they are deemedare applicable for other types of volumetric anomalies, such as global iceberg. is their higher longevities.

C. Modules:

Rule Synthesizer:
The sockets are programmable rule-matching units that are optimized to match the programmed rule and count the size of Data-Engine are collected by the associated logic and communicated to higher control and analysis layers forfurther processing. BURAQ's data plane is a custom architecture on an FPGA unit that combines a rule processing unit, the Data-Engine, with associated synchronization and data transfer logic. The Data-Engine itself is composed of a number of parallel rule to a systolic array.

D. Response Evaluation:

It is through the response-evaluation engine that the discussed streaming algorithms provide the intermediate results from the back end in assisting the dynamic synthesis of intermediatethat translates the user requirements into socket-deployable the data plane and a Response-Evaluation Engine in automating the exploration of the vast is where the user programs the high-level formulation of the measurement requirements.

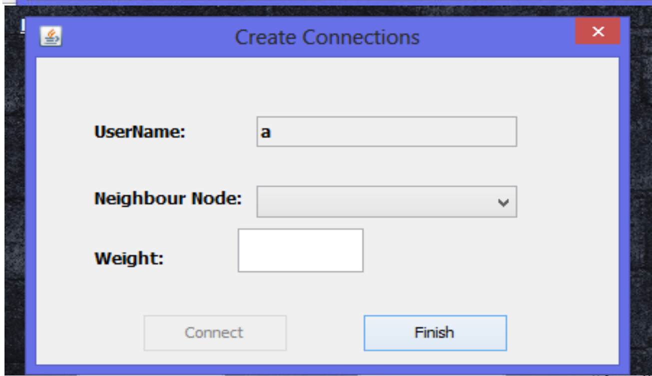
E. Anomaly Detection:

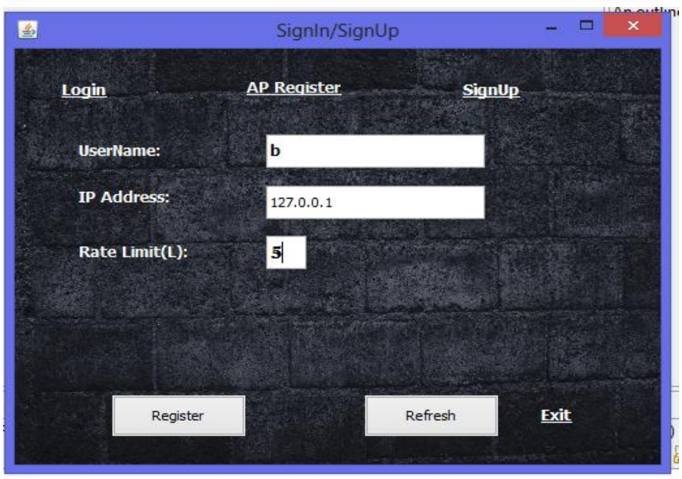
The BURAQ platform uses PDR to maximize the FPGA utilization in increasing the number of rule-processing sockets. The bounds provide tuning knobs for a user to fine-tune the system's accuracy and latency for any given platform and network conditions. The synthesis of rules, deployment, and collection of their results from the computing platform involves finite latencies, during which streaming packets may miss observation. We aggregate the above latencies together as reprogramming latency. We detach the rule-sets in discrete measurement cycles by assuming the controller only adapts new rules once a given rule-set is completely processed.

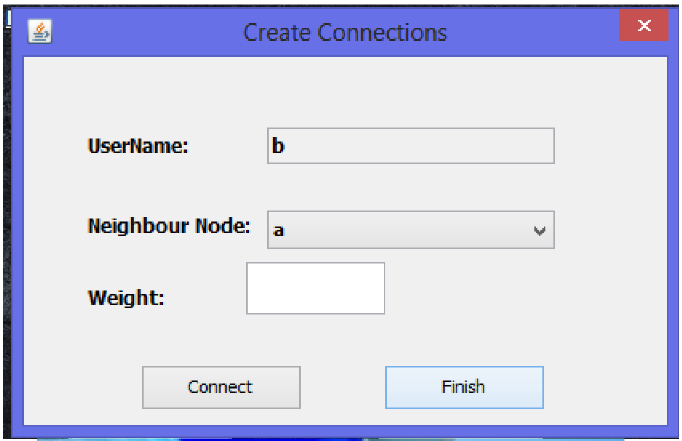
F. Performance Analysis:

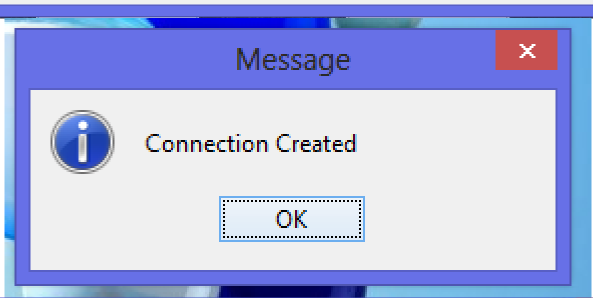
Software solutions face two additional overheads: 1) increased latencies in forwarding packets from a streaming network to the point of measurement (in cases where measurement framework is the software stack where rule processing takes place. The exact latencies are a function of network layout, CPU speed, operating system, and software stack.

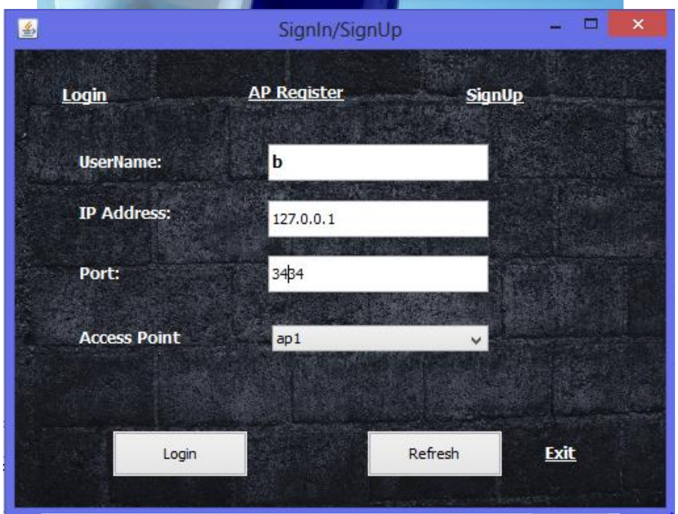
IV. RESULTS

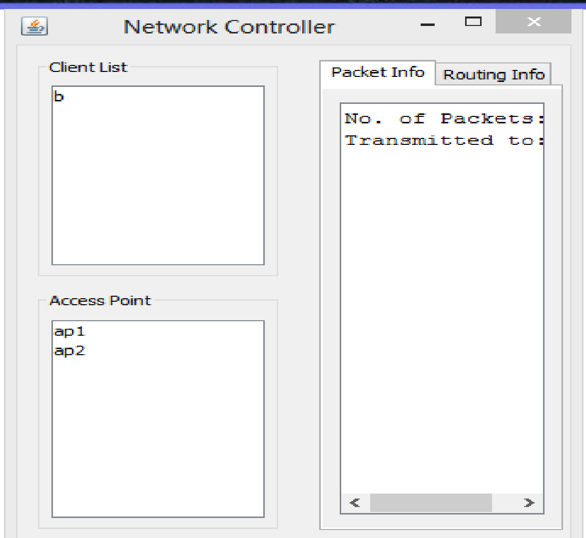


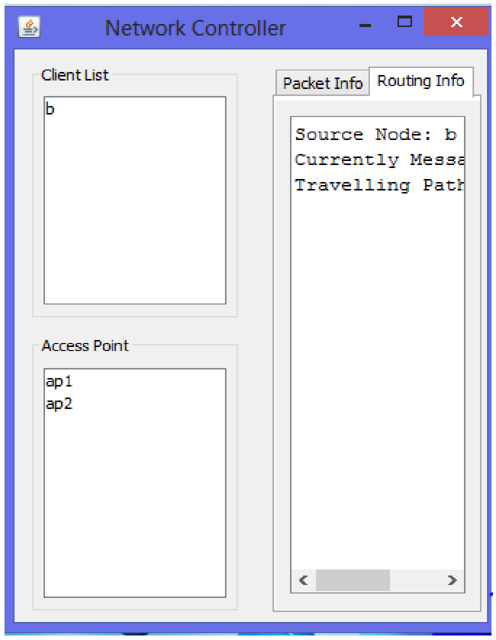












V. CONCLUSION

Traditional are heavily tied with smart algorithms that can by the constraints or opportunities are from measurement platforms. Algorithms l and response. Hardware and software-based measurement as distributed marked 100% improvement in detection accuracy computational complexities.

REFERENCES

[1] Cisco, San Jose, CA, USA,
[2] F. Khan, S. Ghiasi, and C.-N. Chuah, “Adynamically reconfigurable system for closed-loop measurements of, 2012, to be published.
[3] F. Khan, N. Hosein, S. Ghiasi, C.-N. Chuah, and P. Sharma, “Streaming solutions for fine-grained network traffic measurements, 2013 [Online]. Available: http://www.ece.ucdavis.edu/cerl/techreports/2013.
[4] N. Brownlee, RFC 2722, 1999 [Online].
[5] N. G. Duffield, “Sampling for passive Internet.

