ILLINOIS TECH

College of Computing

Integration of GraphBLAS on FABRIC for Network Traffic Analysis

Vaneshi Ramdhony, Hyunsuk Bang, Nik Sultana

Introduction

Analyzing network traffic is crucial for threat detection. This project is the first deployment of GraphBLAS on FABRIC. GraphBLAS uses linear algebra for network traffic analysis, and FABRIC is an international network testbed platform used for research and experimentation in networking.

GraphBLAS forms a network traffic matrix with rows denoting source addresses, columns denoting destination addresses, and values indicating the number of packets between a given source and destination. This integration of GraphBLAS and FABRIC can lead to enhanced scalability and real-time threat response.

Motivation

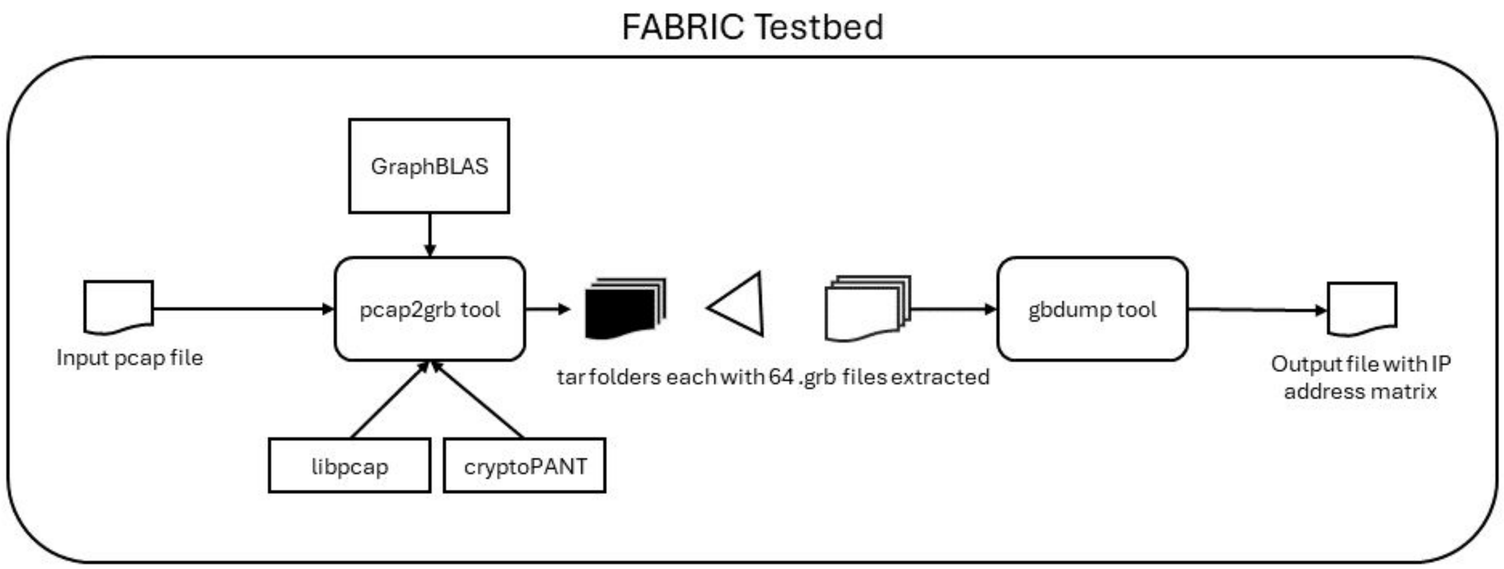
If applied more widely on FABRIC, GraphBLAS can aid with threat identification and potentially enable real-time threat response. By correlating network traffic with contextual data we can gain valuable insights into incidents, which in turn could aid with effective responses. Moreover, the utilization of GraphBLAS for efficient processing of large-scale matrices offers scalability for analyzing network traffic data, making it a potentially suitable option for managing high-speed networks and large testbeds.

Results

- The verified that the same matrix values are obtained as the output files, both locally and on FABRIC. The matrix contains IP addresses, with each row having the following format:

  <source IP addr> <destination IP addr> <num of packets>

- The tools can work with different sizes of input pcap files.



- ❏ pcap2grb: reads packets from a pcap file and saves them into serialized GraphBLAS matrices.
- ❏ cryptoPANT: supports anonymization and de-anonymization of IPv4, IPv6, and MAC addresses.
- ❏ gbdump: dumps the contents of a GraphBLAS matrix to a file/standard output.

Approach

Local

- We installed GraphBLAS and its associated libraries.
- An input pcap file is processed using the pcap2grb tool. Output tar folders are obtained, each containing 64 .grb files.
- The tar folders are extracted and passed as input to gbdump. A script is created to automate running gbdump on each .grb file in all the folders. A matrix containing the IP addresses is generated as the output file.

FABRIC

- We created a Jupyter notebook where GraphBLAS and its associated libraries are installed on FABRIC.
- The input pcap file used on the local testbed is copied to FABRIC to test on the same input.

Validation

- For a single input pcap file, the output files generated on the FABRIC testbed are compared with those generated locally.
- The process is repeated for different pcap files.

Future Work

- Determine how much traffic GraphBLAS can handle.
- Visualize the analysis done by GraphBLAS.

Acknowledgement

We thank Michael Jones and Jeremy Kepner at MIT Lincoln Lab for help with GraphBLAS.