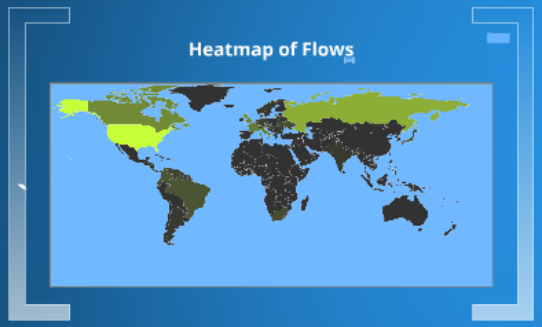
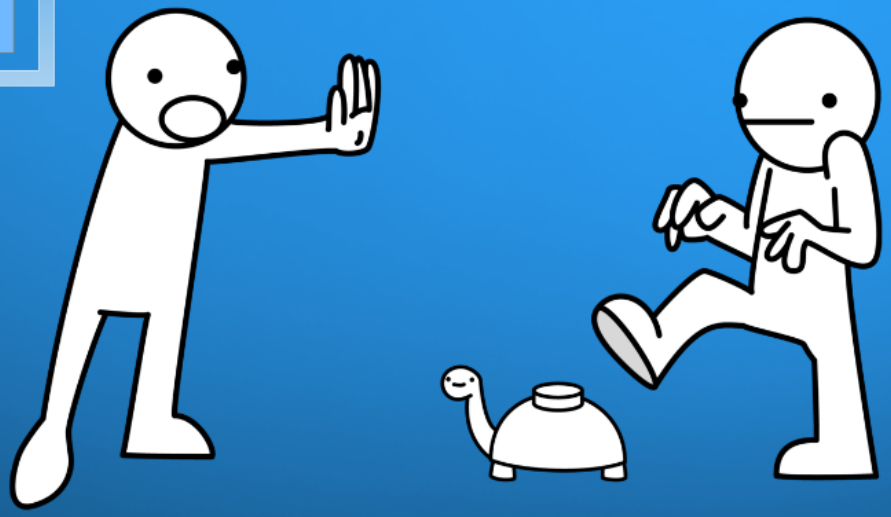
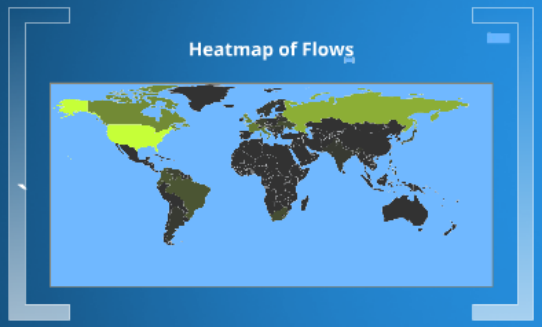
An Exploration of Network Flows for Geolocation and Traffic Visualization to aid in Cyber Defense



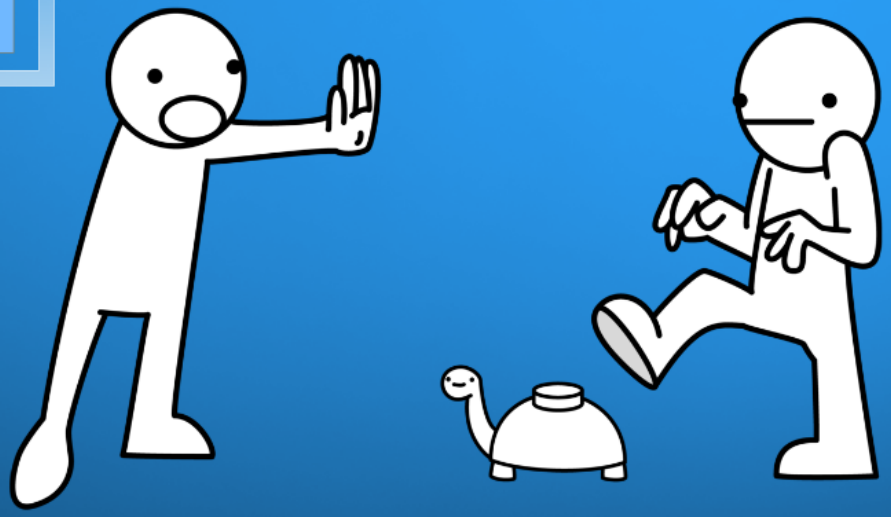
By Sean Pennefather



An Exploration of Network Flows for Geolocation and Traffic Visualization to aid in Cyber Defense



By Sean Pennefather



Research Goals

- An investigation of protocols used to export raw flows.
- Investigation into using raw flows for geolocation and data visualization.
- Investigation into the feasibility of a real-time Geolocation system using network flows to aid in cyber defense.

Approach

- Develop a prototype system capable of collecting and processing raw flows.
  - Process exported flows.
  - Perform geolocation on destination IPs.
  - Render heatmaps for collected flow characteristics.
  - Replay geolocated traffic.
  - Produce a sample monthly report on traffic statistics.

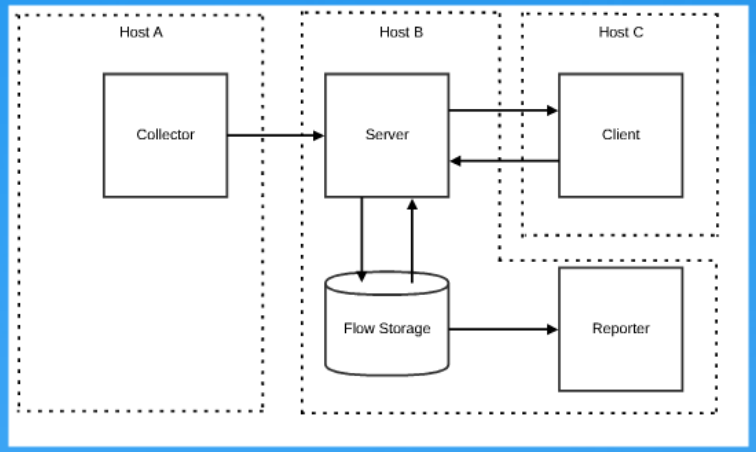
So.. Network Flows

- What are they?
  - A way of representing a stream of unidirectional packets as a single entity.
- The good
  - Less data (less memory and processing needs)
- The bad
  - Less data (less resolution of traffic)
- NetFlow v5, v9
- IPFIX

Scope

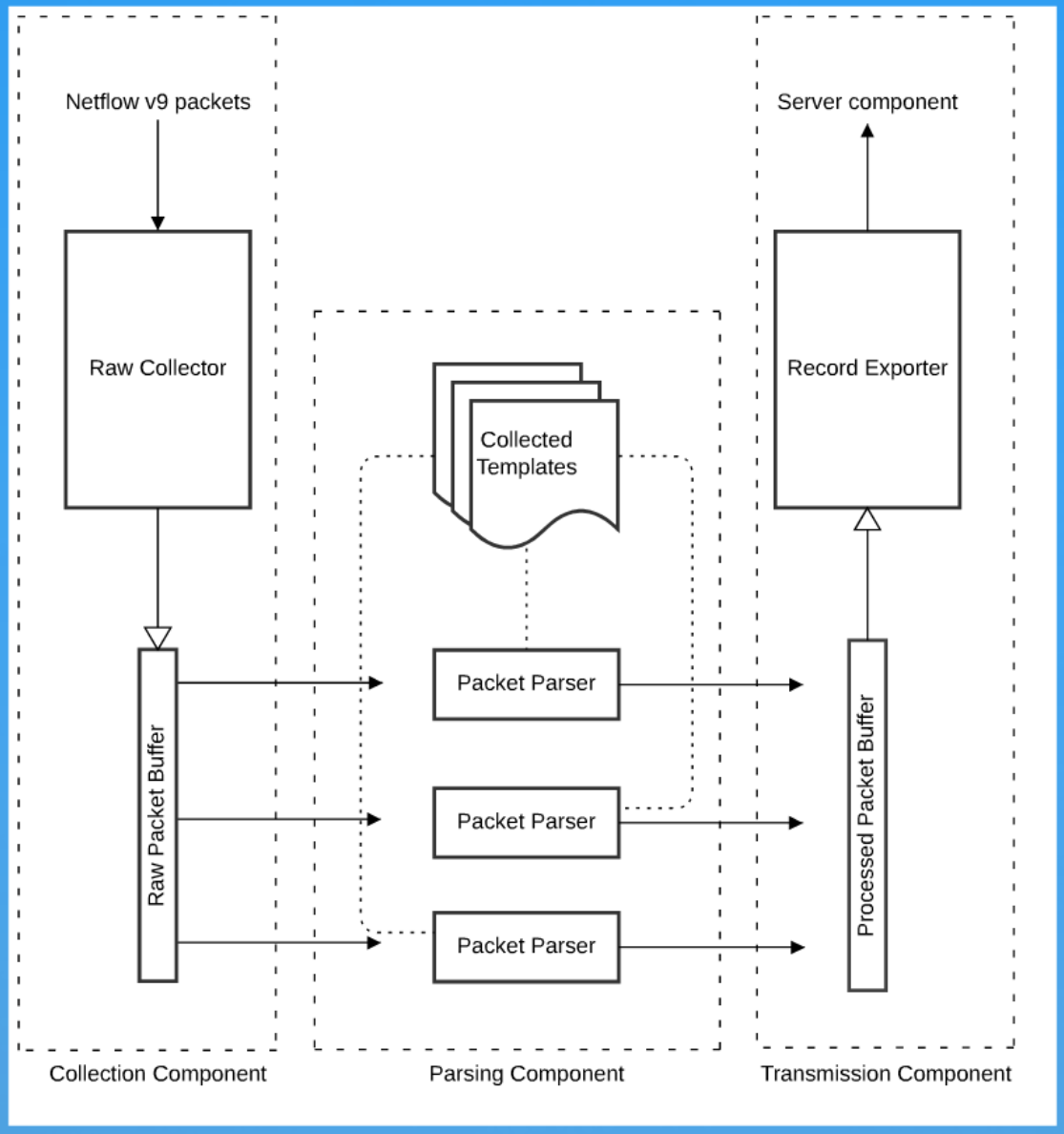
- Implemented to accept NetFlow version 9.
- IPv4 only
- Ignoring options templates and records
- Host based only
- Report is only a sample to show what can be done
- Not overly concerned with memory impact of system.

System Overview



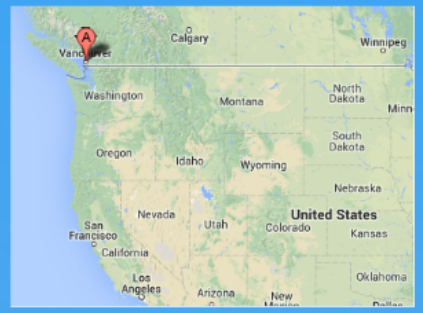
- Four components:
  - Collector: Collects and parses raw flow packets
  - Server: Geolocates IPs, stores records, grunt work
  - Client: User interface of system, renders maps
  - Reporter: Generates a sample report

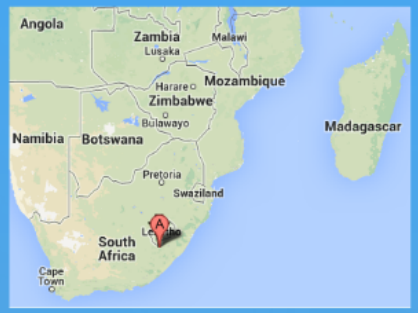
Collector



Geolocation Testing

Google Maps





System

System Timings

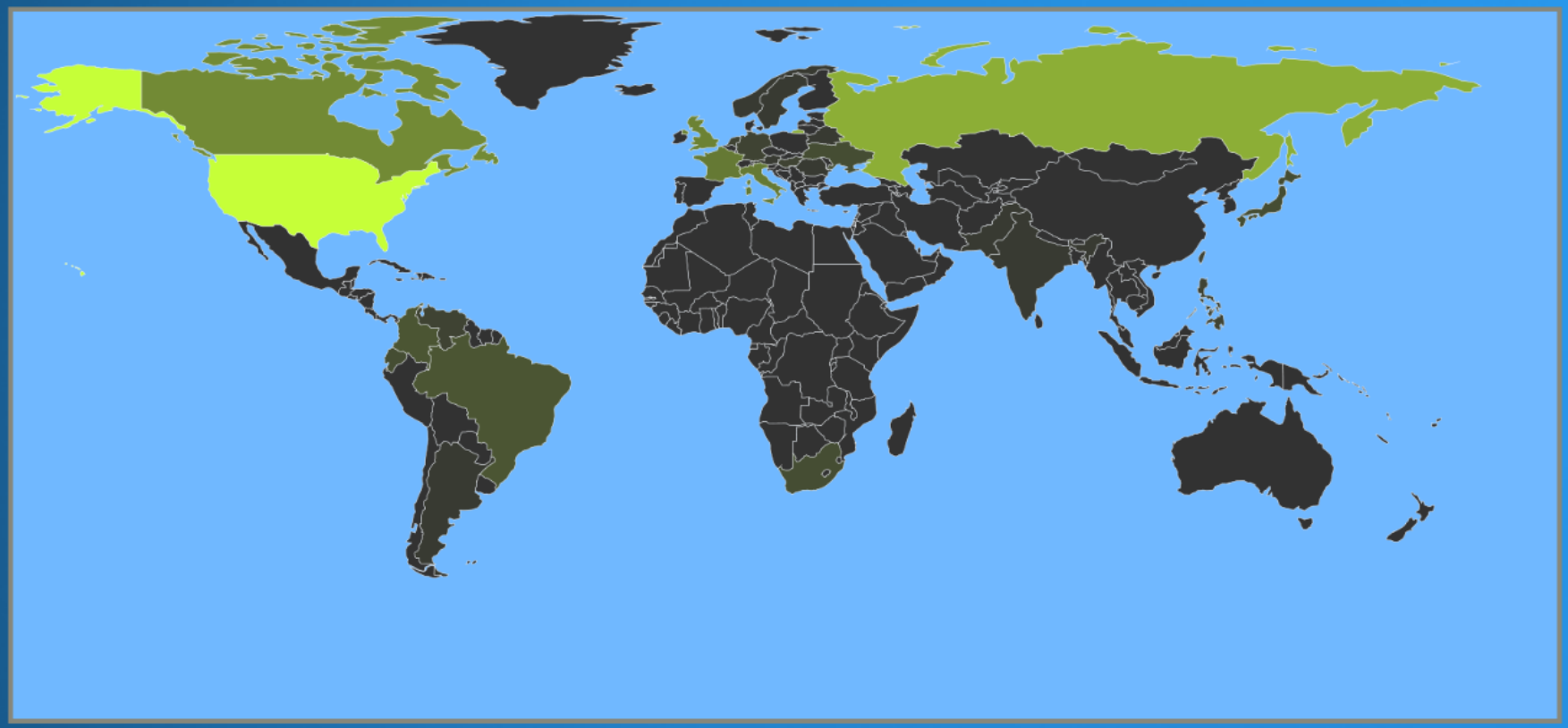
Table 5.3: Sample recorded results of time taken to process a network flow

Test No.	Time Packet Received	Time First Record Sent	Duration [s]
1	1381000626.1056400	1381000626.1600400	0.0543940
2	1381001156.9537300	1381001157.0111500	0.0574150
3	1381001211.8136500	1381001211.8485900	0.0349381
4	1381001301.8416500	1381001301.8949600	0.0533080
5	1381001356.3536600	1381001356.4073800	0.0537219
6	1381001536.2776300	1381001536.3262600	0.0486290
7	1381001686.5536400	1381001686.6038200	0.0501890
8	1381001771.8176900	1381001771.8710300	0.0533390
9	1381001866.4496300	1381001866.4996200	0.0499859
10	1381001921.6336400	1381001921.6834500	0.0498040

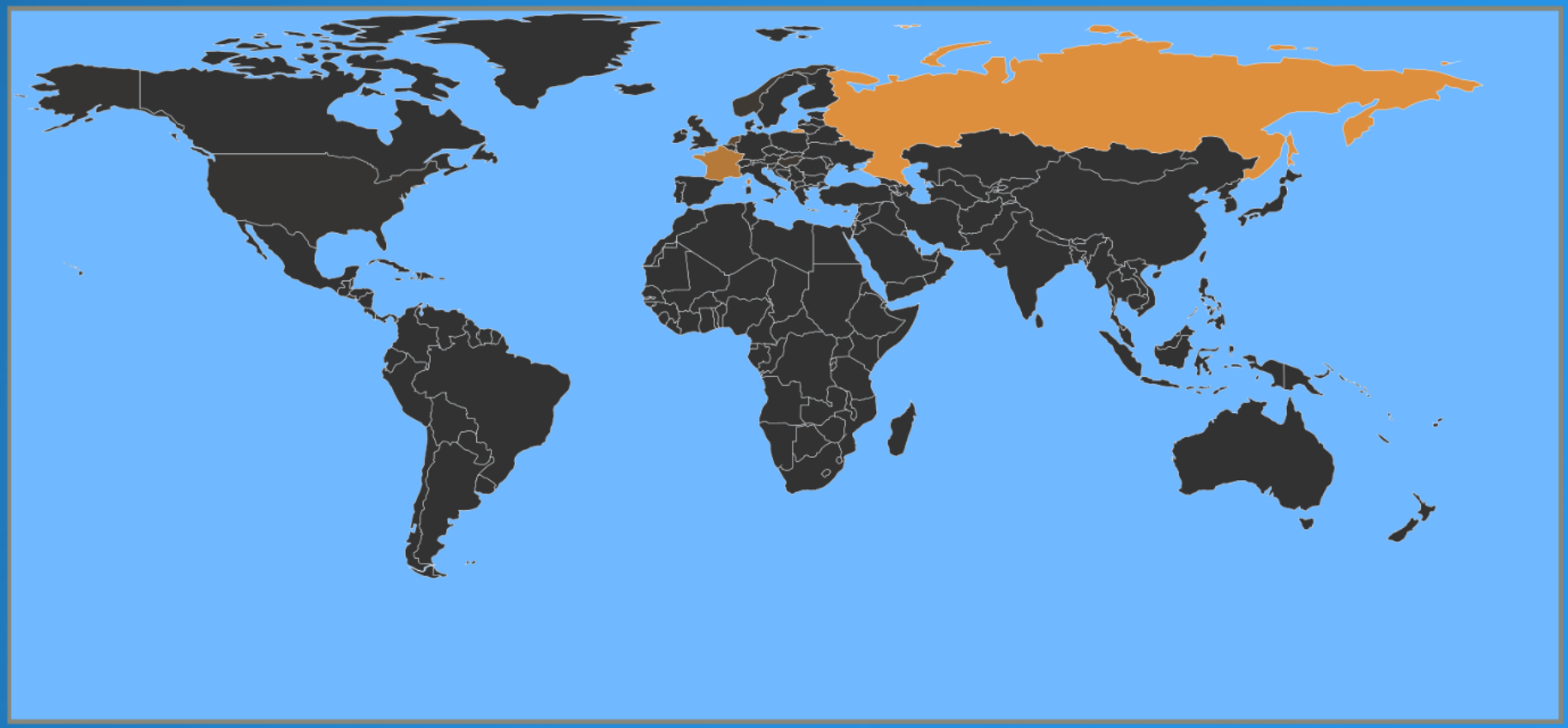
Next, a System Demonstration

Realtime Map

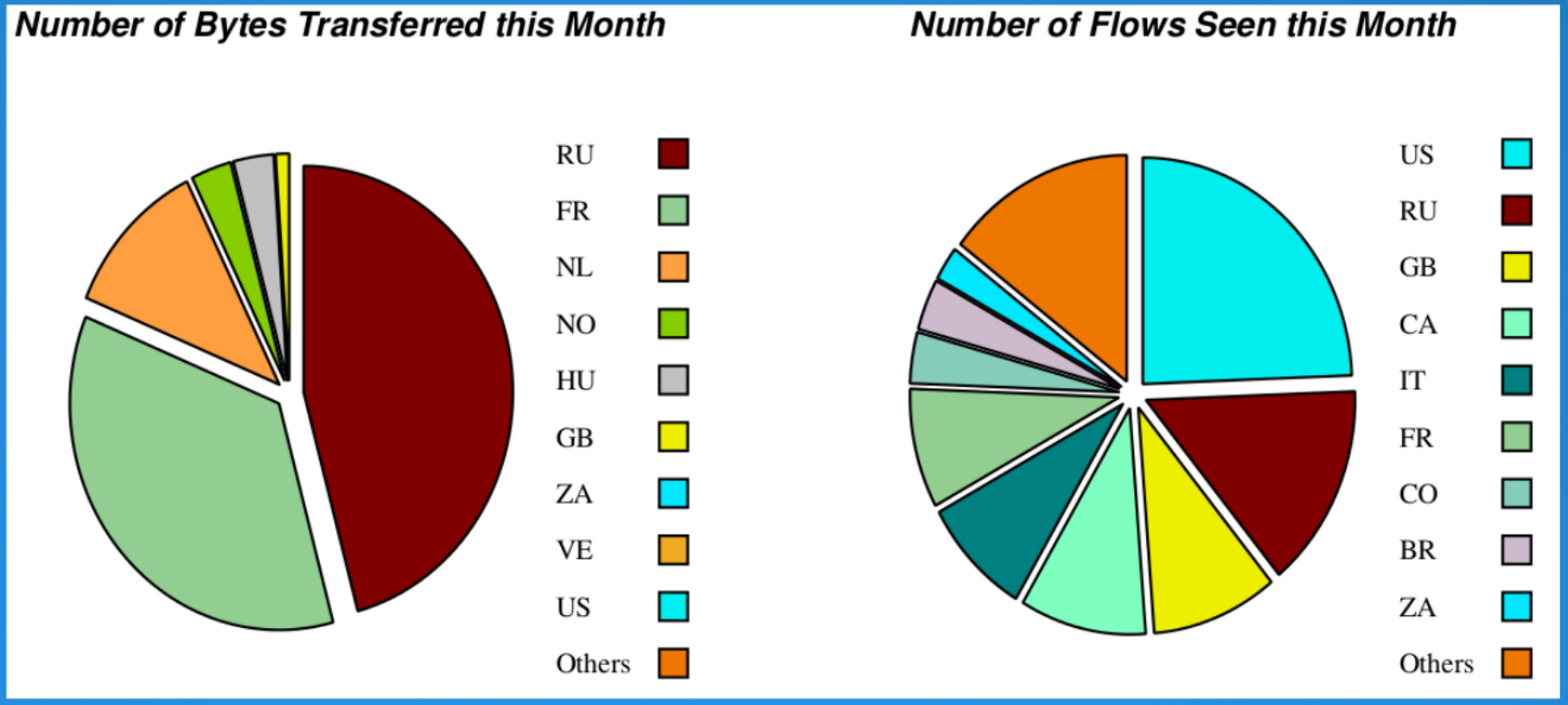
Heatmap of Flows



Heatmap of Bytes



Or another way to look at it...



Conclusions

- Was able to successfully produce a prototype system capable of meeting the defined goals.
  - Geolocation and heatmaps proved successful.
  - Sample report would need improving on...
- Real-time is feasible but requires the collector maintain flow state.

Future Work

- Expand the current system
  - IPv6
  - Better reporting tools
  - IPFIX
- Crteate a better software based flow export system.
  - Handle packet timings correctly
  - Allow for characteristic specification
  - Allow IPFIX export
  - Allows users to define their own characteristics to export.

Questions?