MoWireless MoProblems: Modular Wireless Survey Systems and the Data Analytics That Love Them

Geoff Horvath, Alsatian Consulting, LLC
Winson Tam, Uplynx LLC



Geoff Horvath

- 13 Year US Army Intelligence Officer
  - 10 years in Special Operations units
- Founder, Owner, Consultant – Alsatian Consulting, LLC
  - https://alsatian.consulting
  - Cybersecurity, Signature Reduction, etc to public sector and private organizations/individuals
- BA, Political Science - Drexel University
- MS, Information Systems Engineering - SANS Institute
- 10 GIAC Certifications
  - GIAC Analyst ID : 177730
- Wireless enthusiast
- 2 German Shepherds
  - 13 Year Old - Belle
  - 3 Year Old - Gunner

UPLYNX_

$ whoami

winson_tam



- Software Developer
- Penetration Tester
- Red Teamer
- Proud Cat Dad
- Infrastructure as Code Fanboy
- US Navy Veteran (JSOC)
- Better Slide Designer than Geoff

Why are you doing this?

- Desire to collect large swaths of information to see where things come from
- Tried all-in-one solutions
  - Form Factor issues
  - Power concerns
  - Number of peripheral devices issues
- Large amount of data stored in multiple files/filetypes

Signal Processing and Reconnaissance Kit – Yellow (SPARK-Y) Box

- Modular Wireless Collection
  - WiFi (2.4GHz, 5GHz, 6Ghz)
  - Bluetooth/BLE
  - Cellular Network Parameters (2G-5G)
  - RTL433-Compatible ISM-Band Signals
- Radios:
  - 13x Alfa AWUS036NEH
  - 13x Alfa AWUS036ACM
  - 1x Alfa AWUS036AXML
  - 5x NooElec NESDR SMArt
- Devices/Software:
  - 3x Raspberry Pi 5
    - Raspberry Pi OS, Kismet, RTC Battery
  - 4x Android Phones
  - G-Net Track Pro, WiGLE
- Case
  - REI Trailgate Gear Bin 60L
  - 2x REI Trailgate Gear Cube 10L
  - 1x REI Trailgate Gear Cube 25L



Why Modular

- Less RAM requirements for lots of devices
  - Multiple channels, antennas, technologies means lots of clients to track
  - Distributing collection reduces RAM requirements to manageable, affordable levles
- No need for live processing
  - Not doing analytics on-the-fly, designed to do initial/follow-on analysis
  - Not a live tracking/collection tool
- Reduces bottleneck / distributes collection
  - Fewer hosts, less RAM considerations reduces device backlog
  - Creates robustness for file integrity, entire wardrive not lost if one file corrupted
- Individually Powered
  - No need to carry around a car battery
- Smaller Files
  - Each pi is focused on one technology/frequency range rather than all the data going into one central file
- Scalable
  - New technologies/more sensors can be added just by building the device, no settings changes needed on already-accepted datatypes
- Disparate Locations
  - Sensors can be spread out over large areas, collecting relevant

PacketGlass

What is PacketGlass?	ELK-based Analytics Platform running in Docker What Winson will talk about
Why is it called PacketGlass	Because ELK-GEO didnt sound as flashy Like looking glass, spy glass, or magnifying glass, used to discover fine details out of larger things.
What does it do?	Takes All the data and makes it centralized without concern for enormous file sizes, processor bottleneck Datasources: WiGLE, Kismet, PCAP, Cellular Log File
Why did you do it like this?	Because again, I didnt need a massively powerful and expensive collection rig.
Does it take alot to run?	Intel Core i7-1165G7 @ 2.8 GHz 64GB Ram 2TB SSD

PacketGlass Overview

Ingesting Distributed Wireless Data (Kismet & WiGle)

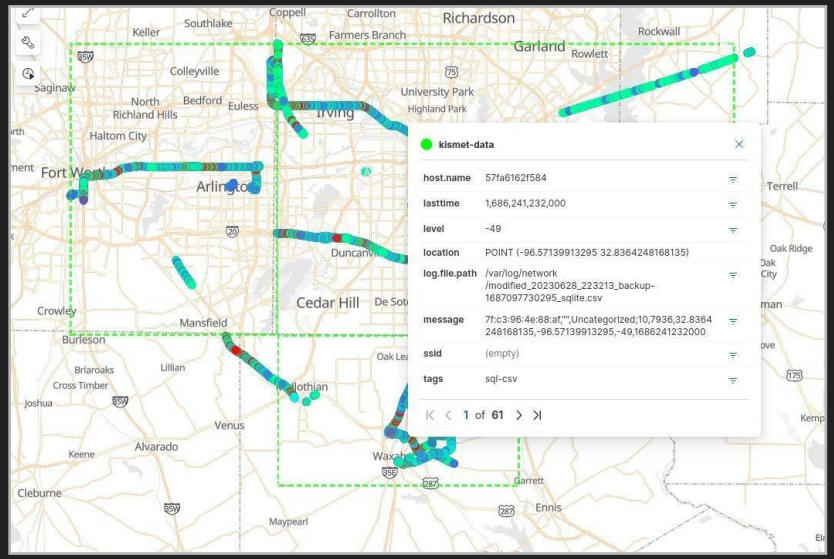
Project Overview

Objectives:

- Dots on Maps
- Easy to use, Low maintenance, Extensible

Requirements:

- Accept .sqlite and .kismet file formats + others
- Sourced from multiple asynchronous devices
- Compatible with Google Drive, SFTP, SCP
- Can be air-gapped



Overview: What is ELK?

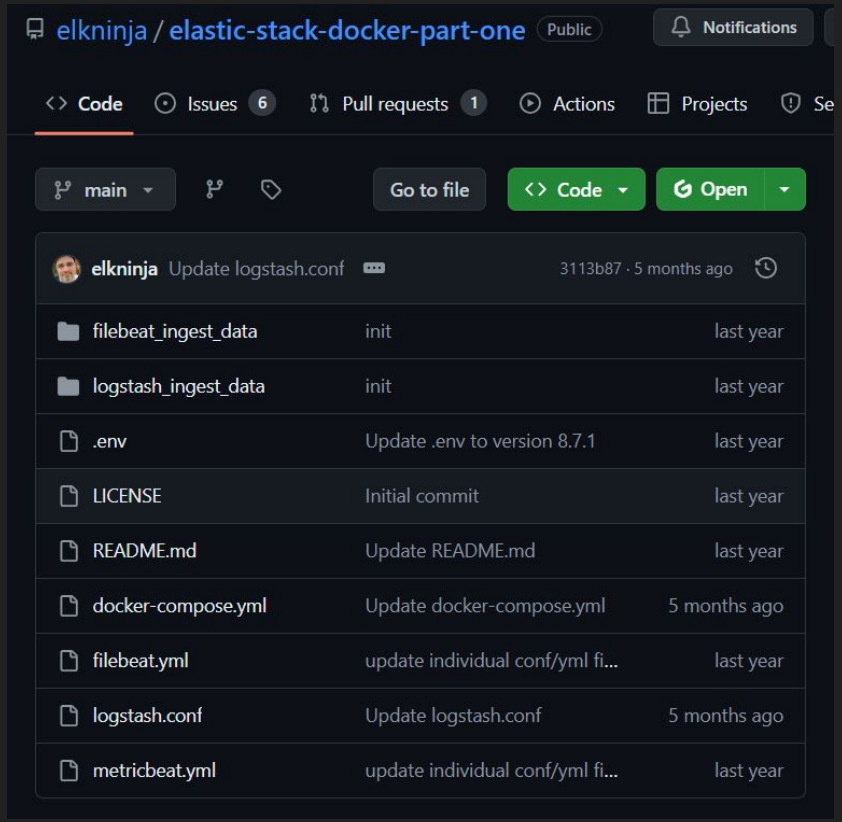
Elasticsearch	Stores, searches, and analyzes data
Logstash	Collects, processes, and routes data from various sources
Kibana	Visualizes and interacts with data through a web interface
—	
(Maptiler)	Stores and serves maps, offline self-hosted possible

Overview: Containerization

- Props to elkninja on github
- Uses `Docker Compose`
- es nodes, logstash, kibana, filebeat, maptiler

—

- Added ${LOG} to .env
- Added maptiler



Overview: Ingestion Workflow

process-data.py monitors *ingest* dir and copies/moves files



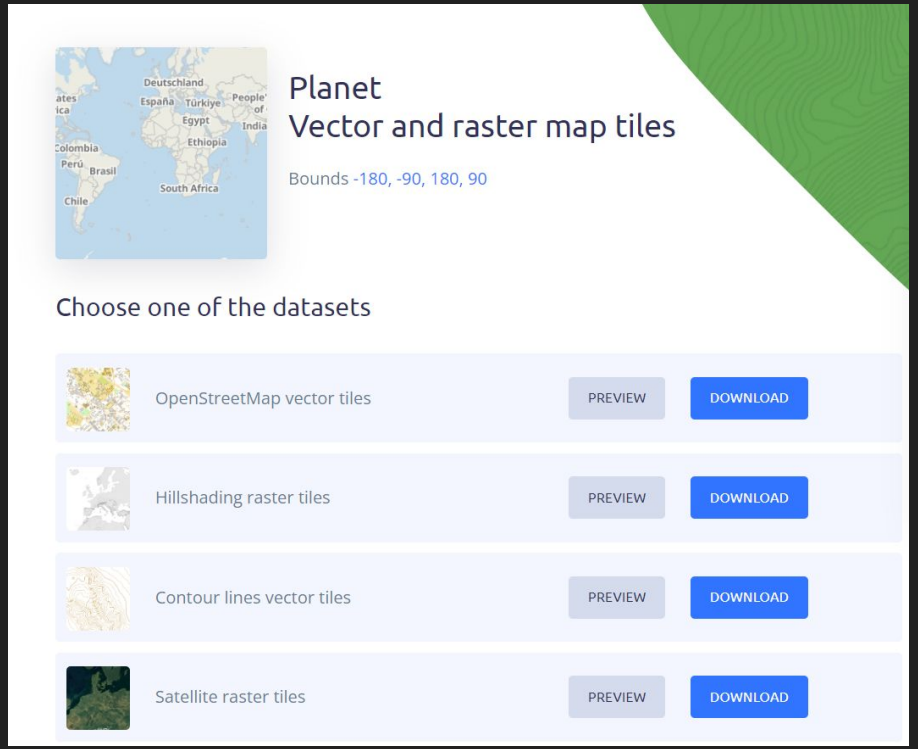
Challenges

- Offline Kibana Maps requires a Maps Server (Elastic Maps Server = $$$)
- Logstash doesn’t natively support .sqlite (WiGLE) nor .kismet files
- Data may contain pcaps or geopoints
- Logstash JSON codec isn’t a fan of multiline/pretty JSON (yes, multiline exists)

Solution: Offline Maps

MapTiler Server

- Self-Hosted Base Maps
- map.tilemap.url:*http://localhost:3650/api/tiles/maptiler-satellite/{z}/{x}/{y}.png*
- Open Street Maps are available



Solution: Wigle (.sqlite) Support

- Use sqlite3 to extract fields of interest and output as .csv

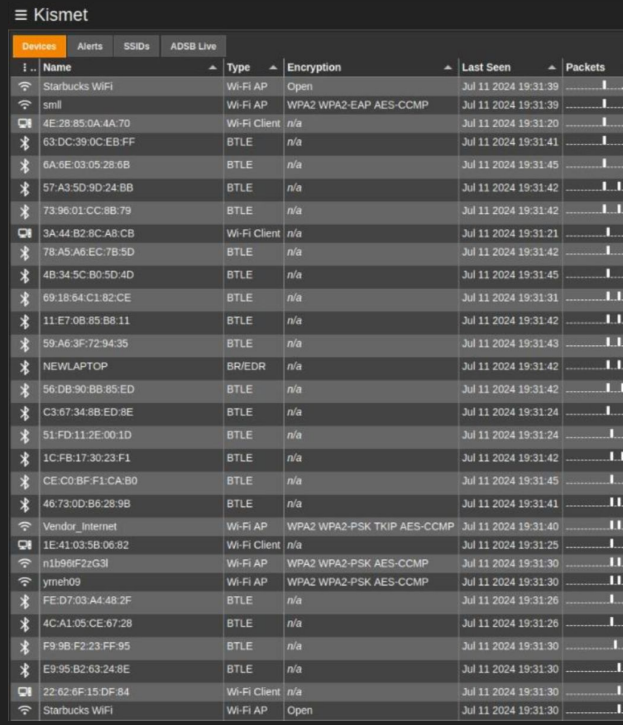
- sqlite3 -header -csv ‘{file}’ ‘select location.bssid, network.ssid,... from network inner join location on location.bssid=network.bssid’ > ‘{file}.csv’

- Create ElasticSearch Indexing Template

```
select
    location.bssid,
    network.ssid,
    network.capabilities,
    network.frequency,
    location.lat,
    location.lon,
    location.level,
    network.lasttime
from
    network
inner join
    location
on
    location.bssid = network.bssid
```

Solution: Kismet (.kismet) Support

- Use kismetdb_dump_devices to extract into a .json file
- Create a ElasticSearch Template for Kismet_Json
- Use kismetdb_to_pcap to extract .pcapng
- Use tshark to convert .pcapng to json (PacketBeat isn't really optimized for this style of ingestion)
- Prettier json to single lines, /n delimited
- Create ElasticSearch indexing template



Analytic Use Cases

- Single Device Identification (demo)
- AP Characterization (demo)
- Area Survey Data (demo)
- Bluetooth (demo)

Special Thanks

Ming + Packet Hacking Village Committee
elkninja
D4rkm4tter
Dragorn
WiGLE Project
Gyokov Solutions
Mark Winters

Questions?