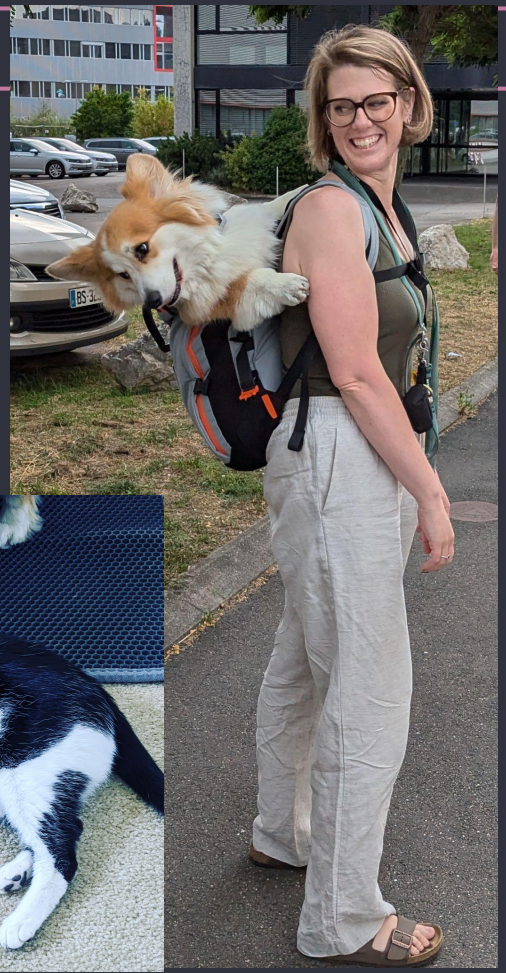
Forensic Flows, but make them better

Jessica Wilson, Brucon 2024

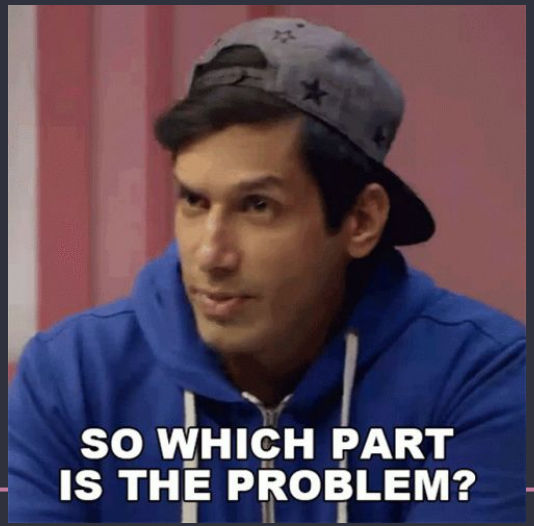
whoami

- Security engineer focused on DFIR
- A nerd
- A corgi and cat mom



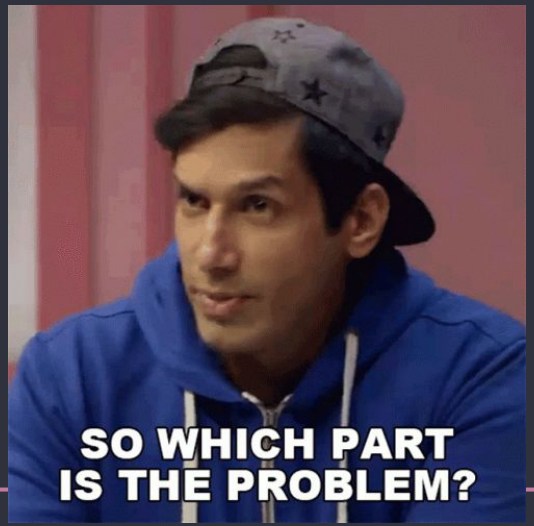
What’s the problem?

- Too few people to be subject matter experts in forensics



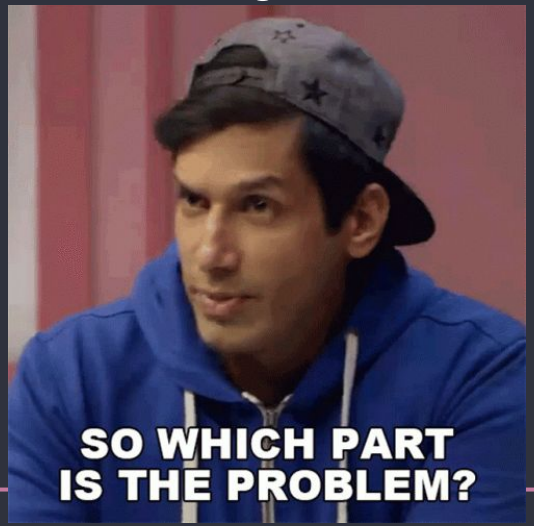
What’s the problem?

- Too few people to be subject matter experts in forensics
- Too many tools that require training



What’s the problem?

- Too few people to be subject matter experts in forensics
- Too many tools that require training
- Not enough automation causing inconsistent investigations



Mission statement

We want all investigators to be empowered to answer any question that arises from an alert within our pipeline with ease. If we simplify forensics, it can be incorporated into everyday investigations

Needs

1. Capacity to grab artifacts on the fly

Needs

1. Capacity to grab artifacts on the fly
2. Ability to process and display evidence

Needs

1. Capacity to grab artifacts on the fly
2. Ability to process and display evidence
3. Scalable environment that runs independently of the investigator

Needs

1. Capacity to grab artifacts on the fly
2. Ability to process and display evidence
3. Scalable environment that runs independently of the investigator
4. Source controlled workflows

Needs

1. Capacity to grab artifacts on the fly
2. Ability to process and display evidence
3. Scalable environment that runs independently of the investigator
4. Source controlled workflows
5. Easy integrations

1. Capacity to grab artifacts on the fly

Define your artifacts

- What do you need to collect?
- Where does that artifact come from?
- If you could have anything, what would it be?

GRR Rapid Response

- Scalable open source live forensics platform
- Supports macOS, Windows, Linux, Kubernetes
- Easy to use API Python client
- Maintained by Google
- Good community support

GRR BETA

Collect from client | Collect from fleet | Use the old UI

g62jw6p239.home | Access granted – *6 days left* | Online

workstation × + Add label

ID:	C.d926ee3d92b34edd	Agent:	grr darwin amd64 3471
FQDN:	g62jw6p239.home	Agent built:	2024-04-29 12:08:35 UTC 4 months ago
OS:	Darwin 14.6.1 23.6.0	First seen:	2024-07-30 09:42:44 UTC 1 month ago
Users:	itadmin, jessicawilson	Last seen:	2024-09-06 09:19:33 UTC 8 minutes ago

View details

All flows | Browse collected files & metadata

Collection

Search for a flow

Common flows

Collect files | Collect browser history | Collect forensic artifacts | Collect path timeline | Interrogate | Osquery

All flows Browse collected files & metadata

Collection

Selected Flow

Collect forensic artifacts

Artifact

Darwin

AllRunningProcessBinaryFiles

Download binaries of all the running processes.

Collects artifact `ListProcessesGrr`

BashShellConfigurationFile

Bourne Again shell (bash) configuration files.

Collects file `%%users.homedir%%/.bash_logout` and 6 more

BashShellHistoryFile

Bourne Again shell (bash) history files.

Collects file `%%users.homedir%%/.bash_history`

Access approval Access granted – *6 days left*

All human flows

Collect files from exact paths
F374BD1D21E498F7

me@jessicawilson.us – 2024-09-06 09:27:34 UTC

1 result

Download files

Collection successful

Flow arguments

/etc/hosts

All flows | Browse collected files & metadata

- /
  - Library
  - System
  - Users
    - itadmin
    - jessicawilson
      - Downloads
      - Library
  - etc
  - private

jessicawilson

List directory | List directory & subdirectories

Name	Size	A-time	M-time	C-time	B-time
.bash_history	83 B	2024-04-23 14:54:24 UTC	2024-04-23 14:54:24 UTC	2024-04-23 14:54:24 UTC	
.zsh_history	177.58 KiB	2024-09-04 14:53:35 UTC	2024-09-04 14:53:35 UTC	2024-09-04 14:53:35 UTC	
Downloads					
Library					

Velociraptor

- Open Source live forensics platform
- Written in Golang
- API through GRPC
- Powered by Rapid7

EDR Vendors

Many EDR vendors offer a collection aspect from their product

- Crowdstrike Real Time Response
- Carbon Black Live Response
- SentinelOne Full Remote Shell

LibCloudForensics

- Open Source library to interact with cloud resources
- Written in Python
- Maintained by Google
- Supports Azure, AWS, and GCP
- Wrapper around API calls to cloud vendors

Any tool that can access your data where it lives

2. Ability to process and display evidence

Turbinia

- Open Source framework for distributed forensic workflows
- Runs
  - Plaso/Log2timeline
  - Docker Explorer
  - Container Explorer
  - Yara
  - BinaryExtractor
  - BulkExtractor
  - And many more
- Written in Python
- Full API server
- Maintained by Google, and good community support

Log2Timeline/Plaso

- Open source framework for automatic creation of timelines
- Written in Python
- Maintained by Google
- Can parse logs such as:
  - Windows Event Logs
  - Browser History
  - FSEvents
  - Cups logs
  - NFTS logs
  - And so many more

Timesketch

- Open source collaborative timeline investigation tool
- Automatically analyze events to highlight critical items
- Written in Python
- Python API client
- Maintained by Google
- Good community support

timesketch
Digital Forensics Timeline Analysis

testing my own report

SHARE

Select an investigative question

*

+ ADD TIMELINE + ADD MANUAL EVENT SELECT ALL UNSELECT ALL

bigquery_queries_timesketch.j...	495	chrome_dlp_timesketch.jsonl	388	command_line_timesketch.jsonl	54
drive_downloads_timesketch.j...	10	externally_sent_email_timeske...	20	notable_events_timesketch.js...	38
related_security_alerts_timesk...	6	screenshots_timesketch.jsonl	102	third_party_storage_lookups_ti...	1
usb_attach_events_timesketc...	29	usb_workstation_write_events...	164	writes_to_downloads_timeske...	442

ADD TIMEFILTER

1-40 of 2073 events (0.045s)

Rows per page: 40 1-40 of 2073 1

Datetime (UTC)	message

1-40 of 1749 events (0.096s)

Rows per page: 40 1-40 of 1749 1

Datetime (UTC)	message	
2023-11-30T14:27:15.000Z	me@jessicawilson.us had notable event: Cloud: interactive shell detected on k...	notable_events_timeske...
2023-11-30T14:53:34.000Z	me@jessicawilson.us had notable event: Cloud: interactive shell detected on k...	notable_events_timeske...
1 days		
2023-12-02T09:04:56.000Z	Username 'jessicawilson" had event 'XarFileWritten' with target file name '/Use...	writes_to_downloads_ti...
2023-12-02T09:04:57.000Z	Username 'jessicawilson" had event 'XarFileWritten' with target file name '/Use...	writes_to_downloads_ti...
2 days		
2023-12-04T09:52:37.000Z	User 'me@jessicawilson.us' sent email with details: subj: Accepted: Snap/Goo...	externally_sent_email_ti...
2023-12-04T13:16:53.000Z	User 'me@jessicawilson.us' sent email with details: subj: Accepted: Lunch at ...	externally_sent_email_ti...
2023-12-04T14:06:17.000Z	Username 'jessicawilson" had event 'PdfFileWritten' with target file name '/Use...	writes_to_downloads_ti...
2023-12-04T14:07:03.000Z	Username 'jessicawilson" had event 'PdfFileWritten' with target file name '/Use...	writes_to_downloads_ti...
2023-12-04T15:21:23.000Z	me@jessicawilson.us had notable event: Cloud: interactive shell detected on k...	notable_events_timeske...

Datetime (UTC)	message
2023-12-02T09:04:56.000Z	bad Username 'jessicawilson" had event 'XarFileWritten' with target file name '...

1 writes_to_downloads_ti...

Size	32313
SourceFileName	
UserName	jessicawilson
data_type	writes_to_downloads
datetime	2023-12-02T09:04:56+00:00
event_platform	Mac
event_simpleName	XarFileWritten
falcon_aid	fd06819b0e6d4725b2dfa14d6cd8763f
filename	/Users/jessicawilson/Downloads/.com.google.Chrome.XqBCkK
	Username 'jessicawilson" had event 'XarFileWritten' with target

Comments

J

me@jessicawilson.us
2024-09-06 09:57 (2 minutes ago)
This is an interesting event!

Add comment

Context search 1S 5S 10S 60S 5M 10M 30M 60M REPLACE SEARCH

Showing context for event:

2023-12-02T09:04:57.000Z Username 'jessicawilson" had event 'XarFileWritten' with target file name '/Users/jessicawilson/Downloads/Unconfirmed

1- of 2 events (0.008s)

Datetime (UTC)	message
2023-12-02T09:04:56.000Z	Username 'jessicawilson" had event 'XarFileWritten' with target file name '/Users/jessicawilson/Dow
2023-12-02T09:04:57.000Z	**Username 'jessicawilson" had event 'XarFileWritten' with target file name '/Users/jessicawilson/Do**

+ SELECT ALL — UNSELECT ALL

Select timelines for analysis

	Name	Description
	Account finder	List accounts detected by the feature extraction analyzer.
	BigQuery matcher	Match pre-defined event fields to data in BigQuery tables
	Browser search terms	Extract search terms from various search providers
	Browser timeframe	Determine user activity hours by finding the frequency of browsing events
	Chain linked events	Chain together events that can be described as linked, either by sharing some common entities, or one event being a derivative of another event
	Domain	Extract domain name from event, tag common and rare domains as well as mark known CDNs
	EVTX gap	Detect gaps in EVTX logs
	Feature Extractions	Runs all feature extraction plugins on selected timelines. Currently implemented extractions: * regex features * winevt features.
	Geolocate IP addresses (MaxMind Database based)	Find the approximate geolocation of an IP address using a MaxMind GeoLite2 database, available from https://maxmind.com
	Geolocate IP addresses (MaxMind Web client based)	Find the approximate geolocation of an IP address using a MaxMind GeoLite2 web client API, available from https://maxmind.com

Cuckoov3 Sandbox

- Open source dynamic malware analysis system
- Primarily Python
- Can detonate a multitude of files

Capev2 Sandbox

- A sandbox is used to execute malicious files in an isolated environment whilst instrumenting their dynamic behaviour and collecting forensic artefacts.
- Derived from Cuckoo v1
- Written in Python
- Primarily for Windows files

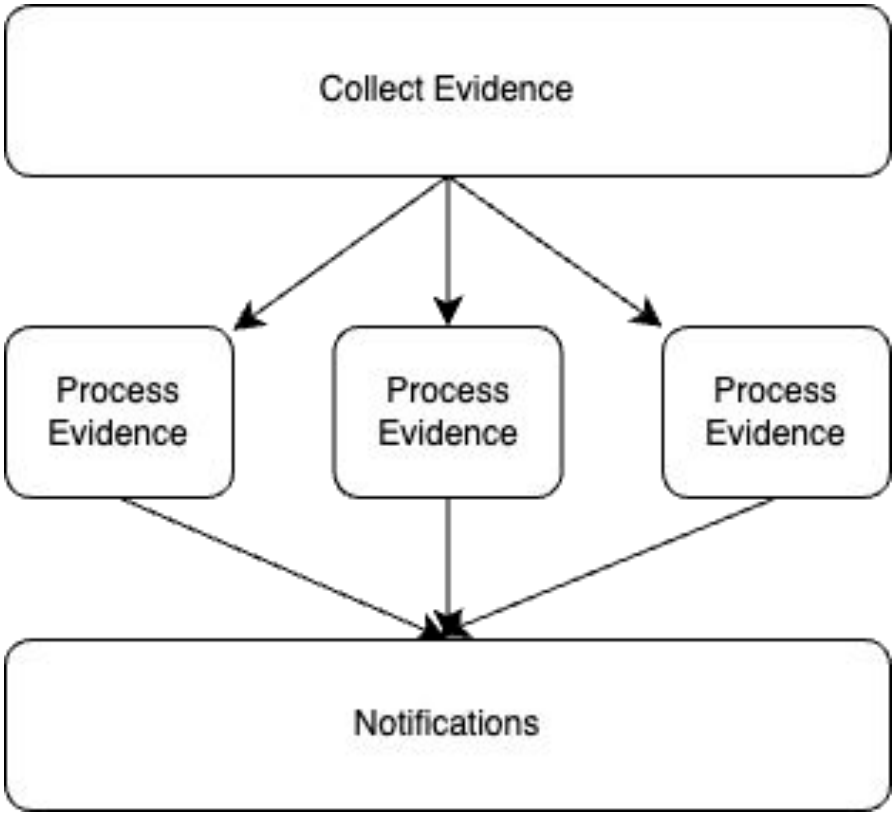
AssemblyLine 4

- A scalable file triage and malware analysis system integrating the cyber security community's best tools
- Open source project maintained by Cyber Center Canada
- Written in Python
- Supports a lot of different file types for Windows, Linux, and a bit of macOS

ASSEMBLYLINE

You are limited by your imagination

Okay cool.. But how do these combine together?


Collect Evidence
Process Evidence
Process Evidence
Process Evidence
Notifications

3. Scalable environment that runs independently of the investigator

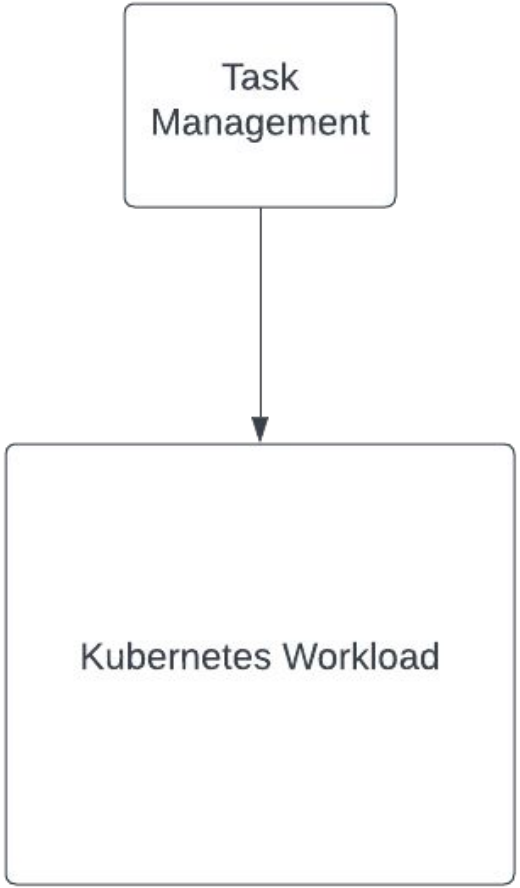
Microservices for the win

- Kubernetes is a great
- Scale your workers per task
- Deployment with Helm
- Source control

Kubernetes Workload

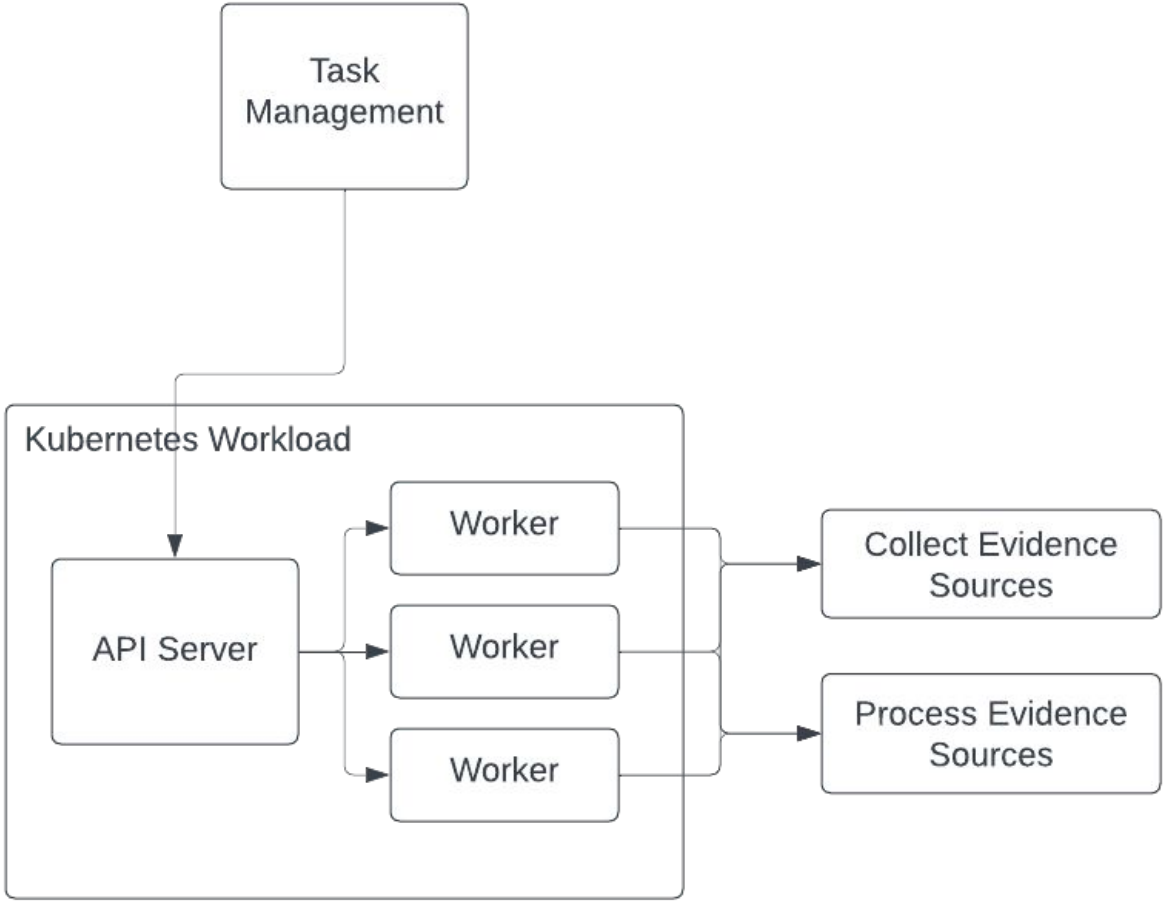
How to run tasks?

- Schedulable
- Scalable
- Easy to interact with
- Solutions:
  - GCP Cloud Tasks
  - Self hosted Celery
  - AWS EventBridge


Task Management
Kubernetes Workload

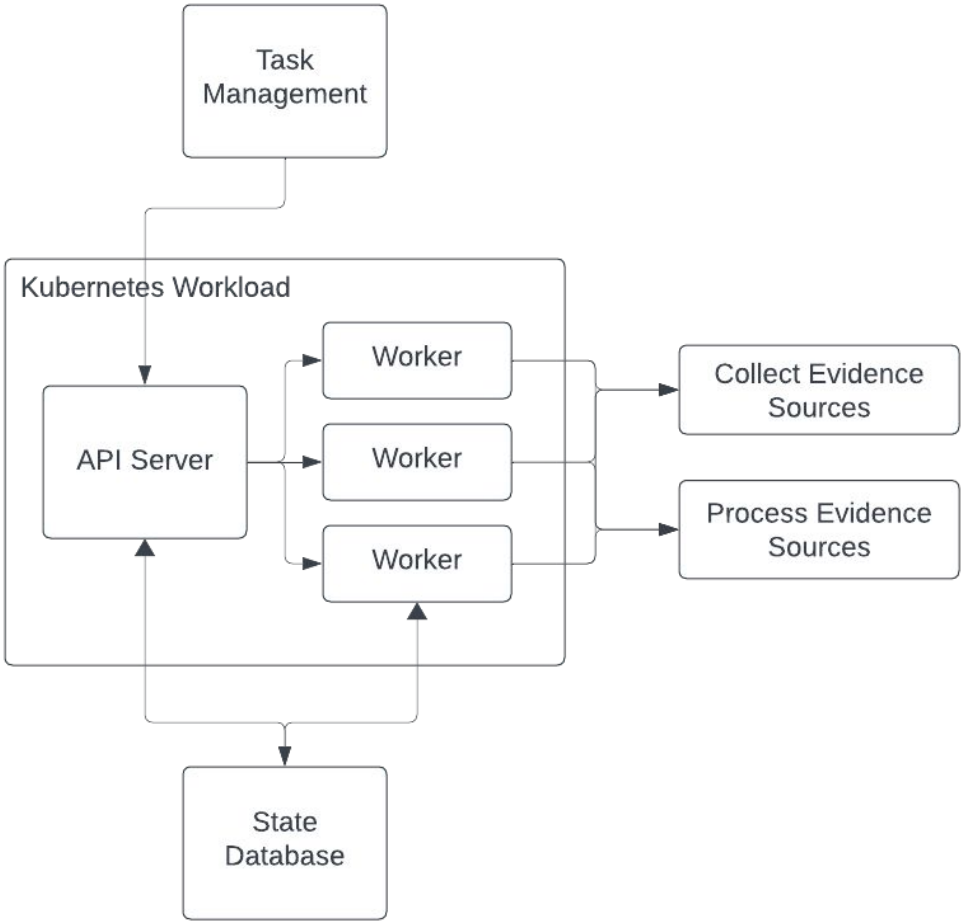
How to scale from the tasks?

- Cloud task queue through HTTP requests
  - Solve with an API server
- Self contained worker
  - Solve with a Kubernetes Job
- Workloads should scale independently


Task Management
Kubernetes Workload
API Server
Worker
Worker
Worker
Collect Evidence Sources
Process Evidence Sources

How to manage state?

- Modular tasks
- State tracking for curious minds
- Auditing is important
- Schemas will change
- NoSQL databases make it easy
  - GCP Cloud Firestore
  - AWS DynamoDB
  - Local MongoDB or another NoSQL db


Task Management
Kubernetes Workload
API Server
Worker
Worker
Worker
Collect Evidence Sources
Process Evidence Sources
State Database

4. Source controlled workflows

Deploy as infrastructure as code

- Consistent environments
- Peer reviewed through pull requests
- CI/CD
- Can use:
  - Terraform
  - CloudFormation
  - OpenTofu
  - Pulumi
  - Whatever you are comfortable with

Kubernetes deployments as code

- Readable deployment charts
- Peer reviewed
- CI/CD
- Can use:
  - Helm
  - Kustomize
  - Carvel
  - Whatever you are comfortable with

Workflows as code

- Single source of truth
- PRs for changes
- Learn more about forensics
- Starting point for deeper investigations

5. Easy integrations

Within your forensic flow system

- Modularize everything
- Normalize field names
- Abstract out methods

Automation

- Normalize your alert data
- API to make requests simple
- Queries per second
- Artifacts can be ephemeral, grab them quickly!

Manual Investigations

- Single pane of glass to use
- One tool to train new investigators on

Now onto the practical portion of the talk

Case study 1: Phishing

Phishing Scenario

- An employee reports receiving a phishing email
- They clicked the email link
- They downloaded and executed the app installer from the link

What next?

Phishing - Old methodology

Understand where that email came from

- Ask the employee and wait for a response
- Check the logs for your email server
- Find the link from the email
- Run that link in a sandbox to understand

Phishing - Old methodology

Grab the malware

- How do you pull it from the machine?
- Is the machine online?
- Does the file still exist on the machine?

Analyze the malware

- Grab the OSINT for the hash
- Run the malware through a sandbox
- Run static analysis on the malware

Phishing - Old methodology

Analyze the machine's behavior

- Pull EDR logs and comb through them
- Pull event logs from the host

Phishing - New methodology

Have 1 button to:

- Submit all links from an email to a sandbox
- Pull a file from a host when available, then route that file to a sandbox
- Pull relevant logs and put them into Timesketch

Asset type: LAPTOP

Assigned to: me@jessicawilson.us

GRR Collection ··· Sandbox Link

Browser Cache (all browsers)

Browser History (all browsers)

Chrome Extensions

Chrome Cache

Chrome History

MacOs Quarantine Events

Collect Files

Connections

Limit (BQ): 100

Last Seen

	Datetime (UTC) ↑	message
	2024-09-07T18:52:35.784Z	https://github.com/rxhanson/Rectangle/releases/download/v0.82/Rectangle0.82.dmg (/Users/jessicawilson/Do

danger_type	15
data_type	chrome:history:file_downloaded
datetime	2024-09-07T18:52:35.784398+00:00
display_name	OS:/tmp/tmp0gzcq6_4/BE06B34F1A419B96_extracted/C.d926ee3d92b34edd_flow_ArtifactCollectorFlow_BE06B34F1A4 Support/Google/Chrome/Profile 3/History
full_path	/Users/jessicawilson/Downloads/Rectangle0.82.dmg
interrupt_reason	0
message	https://github.com/rxhanson/Rectangle/releases/download/v0.82/Rectangle0.82.dmg (/Users/jessicawilson/Downloads, Interrupt Reason: No Interrupt - Success. Danger Type: Safe, Deep Scanned - Download deep scanning identified no proble
offset	802
opened	0
path_spec	{"__type__": "PathSpec", "location": "/tmp/tmp0gzcq6_4/BE06B34F1A419B96_extracted/C.d926ee3d92b34edd_flow_ArtifactCollectorFlow_BE06B34F1A419 Support/Google/Chrome/Profile 3/History", "type_indicator": "OS"}

Case Study 2: Compromised Employee Account

Compromised account scenario

You received an alert for weird behavior from an account. The alert is designed to detect account compromise

What next?

Compromised Account - Old methodology

Determining what the account did

- What IP addresses did this account come from? Are any atypical?
- Was there activity outside of normal working hours?
- Where do all your logs live for this?
- What queries do you need to craft to answer this?

Compromised Account - Old methodology

Determine if and where the credentials were used

- How many systems could the potential attacker have touched?
- Were other accounts affected?
- Legitimate activity vs threat actor activity?

Compromised Account - New methodology

- When the alert is generated
  - All logs are pulled automatically
  - Put into a Timesketch
  - Posted onto the ticket
- Run premade timesketch analyzers
  - Run sigma rules
  - Enrich and tag events

Top 10 sourceIp values

Term	Count
213.206.130.146	14,900
188.60.191.136	104
176.127.123.114	91
144.232.179.30	90

Rare `sourceIp` values

Term	Count
107.178.194.160	1
213.19.204.194	1
2a02:1210:6cd8:6700:cd06:3c1e:49a9:6c0f	1
34.98.143.225	1
107.178.194.224	3
107.178.194.233	3
2a02:1210:6cd8:6700:f414:8223:c415:1098	3
35.187.132.235	4
2a02:1210:6cd8:6700:5ddf:f79d:e506:bb99	5
2a02:1210:6cd8:6700:60b7:ad57:5ce2:3b50	5

	2024-02-22T09:23:38.000Z	me@jessicawilson.us logged in from source ip 213.206.130.146	Logins
	2024-02-22T09:23:38.000Z	me@jessicawilson.us logged in from source ip 213.206.130.146	Logins
	2024-02-22T14:03:56.000Z	me@jessicawilson.us Chrome FILE_DOWNLOAD for /Users/jessicawilson/Downloads/170...	chrome_dlp_timesketch....
	2024-02-22T14:18:49.000Z	me@jessicawilson.us Chrome FILE_DOWNLOAD for /Users/jessicawilson/Downloads/567...	chrome_dlp_timesketch....
	2024-02-22T14:18:59.000Z	me@jessicawilson.us Chrome FILE_DOWNLOAD for /Users/jessicawilson/Downloads/wor...	chrome_dlp_timesketch....
	2024-02-22T14:19:00.000Z	me@jessicawilson.us Chrome FILE_DOWNLOAD for /Users/jessicawilson/Downloads/wor...	chrome_dlp_timesketch....
	2024-02-22T14:19:00.000Z	me@jessicawilson.us Chrome FILE_DOWNLOAD for /Users/jessicawilson/Downloads/wor...	chrome_dlp_timesketch....
	2024-02-22T14:19:15.000Z	me@jessicawilson.us Chrome FILE_UPLOAD for /Users/jessicawilson/Downloads/workflo...	chrome_dlp_timesketch....
	2024-02-23T09:38:17.000Z	me@jessicawilson.us logged in from source ip 213.206.130.146	Logins
	2024-02-23T09:38:17.000Z	me@jessicawilson.us logged in from source ip 213.206.130.146	Logins
	2024-02-23T11:41:57.000Z	User 'me@jessicawilson.us' queried from project 'super-secret-data'.	bigquery_queries_times...
	2024-02-23T11:42:29.000Z	User 'me@jessicawilson.us' queried from project 'super-secret-data'.	bigquery_queries_times...

Case Study 3: Compromised Kubernetes Node

Compromised K8s Node Scenario

You receive an alert from your Cloud Provider that one of your kubernetes nodes was reaching out to a suspicious domain

What next?

Compromised K8s - Old Methodology

Grab the logs

- Can you get network logs?
- Can you get the process logs?

Grab the disk

- Permissions to access the disk?
- Does it still exist?
- Can your team perform dead box forensics to triage?

Compromised K8s - New Methodology

When the alert fires

- Pull relevant time logs for
  - Network
  - Process
  - Container Deployment
- Create a disk image
- Process the disk image with Turbinia to pull out triage artifacts

Case Study 4: Vulnerability Management

Vuln Management Scenario

You have an ask from your vulnerability management team to understand if any Cloud Virtual Machine contains a specific vulnerability.

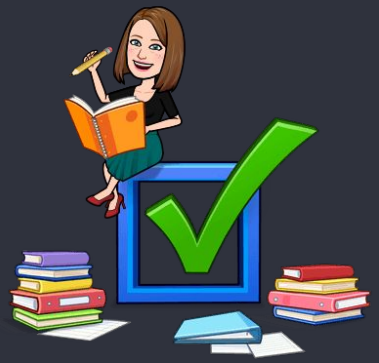
What next?

Vuln Management - Old Methodology

- Pay a company a **lot of money** to do the scanning for you
- Figure out if it’s possible to ssh into every box and scan it

Vuln Management - New Methodology

- Use your forensic capabilities to image disks
- Automate mounting those disks to a premade scanner of your choosing
- Output the results of the scan to your logging system
- Alert on those logs
- Enjoy the benefits of all the historical data from your scanner for your incident response purposes



Lessons Learned

Determine notification paths early

Give clear error messages to investigators

Have metrics, and have them early

Audit trails are best trails

Integration Testing saves headaches

A forensic system is not a replacement for trained investigators

Thank you

Contact me:

Linkedin: https://www.linkedin.com/in/jawilson0502/
Email: me@jessicawilson.us

Any questions?