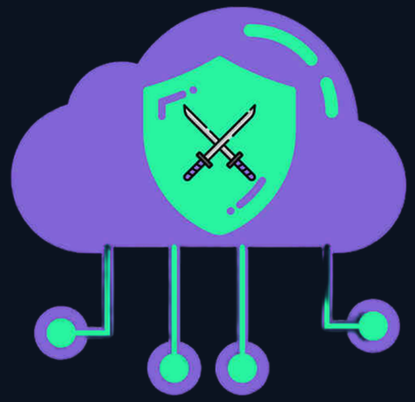
Sweet Deception:
Designing Effective M365 Honey Tokens

Ryan O’Donnell
August 09, 2025



CLOUD VILLAGE

DEFCON

Agenda

1. Background
2. Design
3. Results
4. Conclusion

whoami

- Ryan O’Donnell
- Senior Security Engineer @Microsoft
- Penetration Testing, Red Teaming, Purple Teaming, DFIR
- OSCP, OSEP, GCFA, GREM

Problem

Florian Roth ⚡

@cyb3rops

We're seeing a clear trend: attackers are bypassing the endpoint entirely. Not just avoiding traditional EDR-monitored systems by pivoting to embedded and edge devices, but now also operating purely in the cloud. No shell, no malware, no persistence on the endpoint. Just an OAuth token and full access to whatever's in the victim's Microsoft 365, Google Workspace, or AWS console.

Florian Roth: @cyb3rops

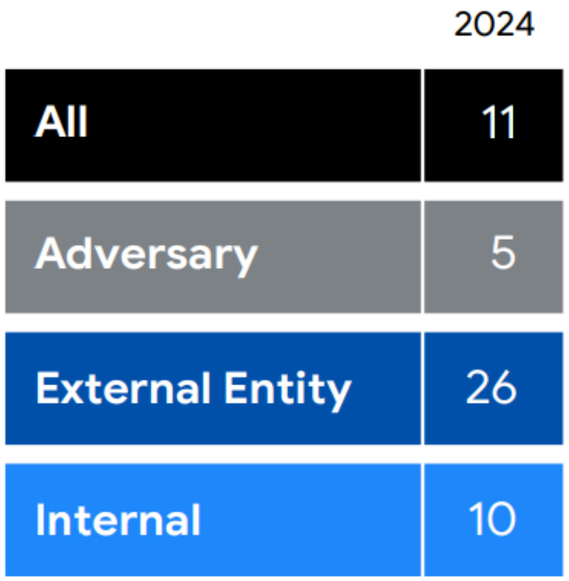
Trend

- Cloud intrusions **increased 136%**
- China based cloud intrusions **up 40%**
- 81% of intrusions were **malware-free**
- Interactive intrusions increased 27%

Crowdstrike 2025 Threat Hunting Report

Time

Median Dwell Time by Detection Source, 2024



The median adversary notification time was just five days, while external partners notified in a median of 26 days. This discrepancy is not surprising given that the vast majority of adversary notifications originate from extortion actors who benefit from monetizing intrusions quickly.

Mandiant M-Trends 2025 Report

Proposal

Canary Tokens can provide **effective** early alerting on attacker **post-compromise** activity in M365.

Background

Cyber Deception

- Deceive attackers using traps, lures, decoys
  - Provoke adversaries **Post-Compromise**
- Goals:
  - **Early Detection + Alerting**
  - Intelligence Gathering
  - Frustration + Delay

Honeypots

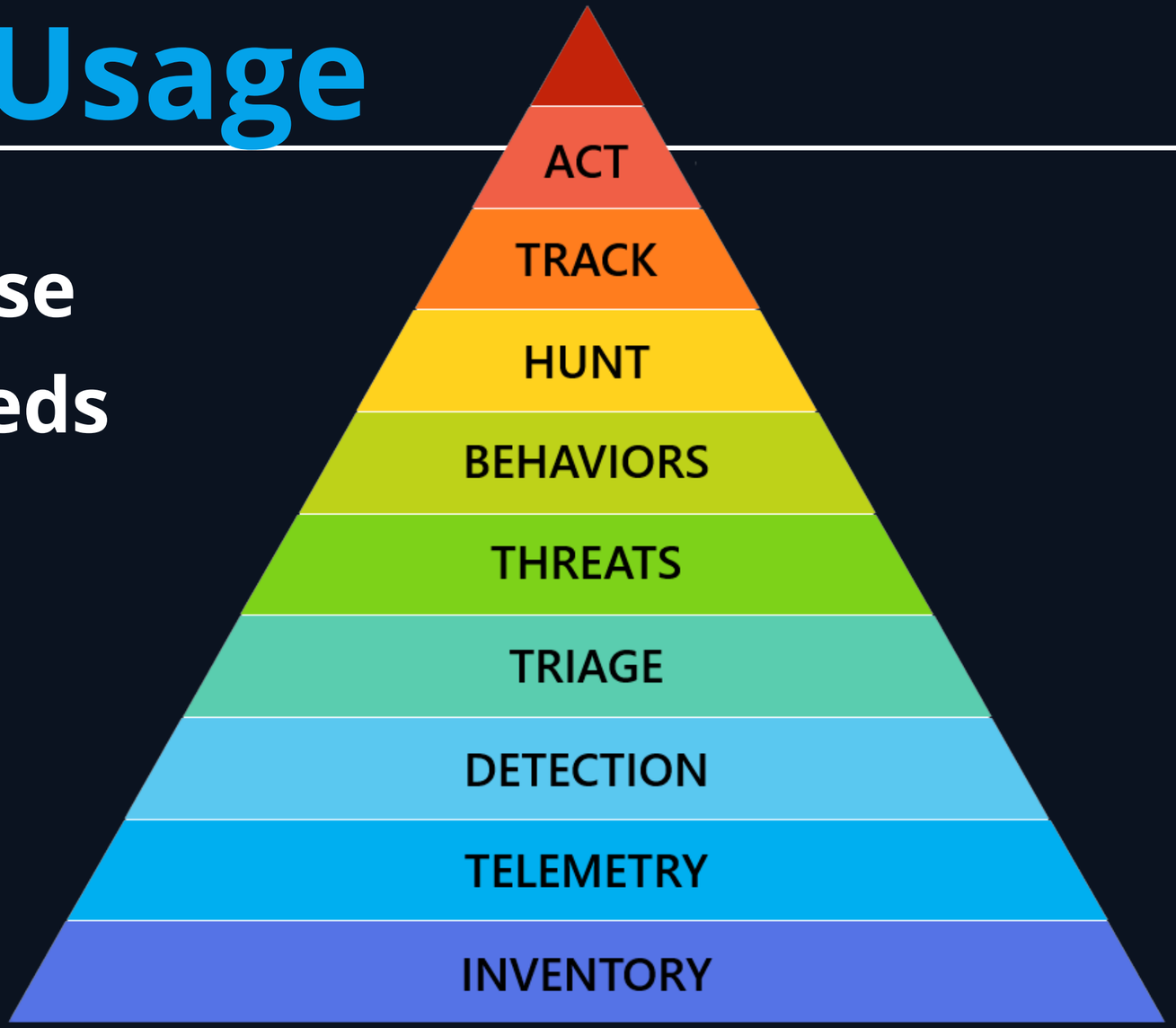
- Decoys designed to attract attackers
- Good source of threat intelligence
- Log behavior

Canary Tokens

- Discrete data elements
  - Credentials , API keys, Docs
- Function as tripwires
- Generate alerts when triggered

Traditional Usage

- **Incident Response Hierarchy of Needs**
  - Matt Swann



Criteria

Attractive

High Fidelity

Non-Invasive

Native

Post-Compromise

Additionally, Microsoft observed Storm-2372 using Microsoft Graph to search through messages of the account they've compromised. The threat actor was using keyword searching to view messages containing words such as username, password, admin, teamviewer, anydesk, credentials, secret, ministry, and gov. Microsoft then observed email exfiltration via Microsoft Graph of the emails found from these searches.

Microsoft Threat Intelligence: Storm-2372

Espionage

After gaining initial access, Void Blizzard abuses legitimate cloud APIs, such as Exchange Online and Microsoft Graph, to enumerate users' mailboxes, including any shared mailboxes, and cloud-hosted files. Once accounts are successfully compromised, the actor likely automates the bulk collection of cloud-hosted data (primarily email and files) and any mailboxes or file shares that the compromised user can access, which can include mailboxes and folders belonging to other users who have granted other users read permissions.

Microsoft Threat Intelligence: Void Blizzard

BEC

Finding a target

The following days after the cookie theft, the attacker accessed finance-related emails and file attachments files every few hours. They also searched for ongoing email threads where payment fraud would be feasible. In addition, the attacker deleted from the compromised account's Inbox folder the original phishing email they sent to hide traces of their initial access.

Microsoft Threat Intelligence: Cookie Theft To BEC

Determination

- Attackers want data!
  - Creds, Keys, Financial Info, RMM tools

Outlook Canary Design

Existing Approach

- **Link** inside email
- Violates the criteria:
  - Notify users – **Non-Invasive**
  - Visible to users – **High Fidelity**
  - 3rd party tool – **Native**
- **Extra steps** for attackers!

Solution

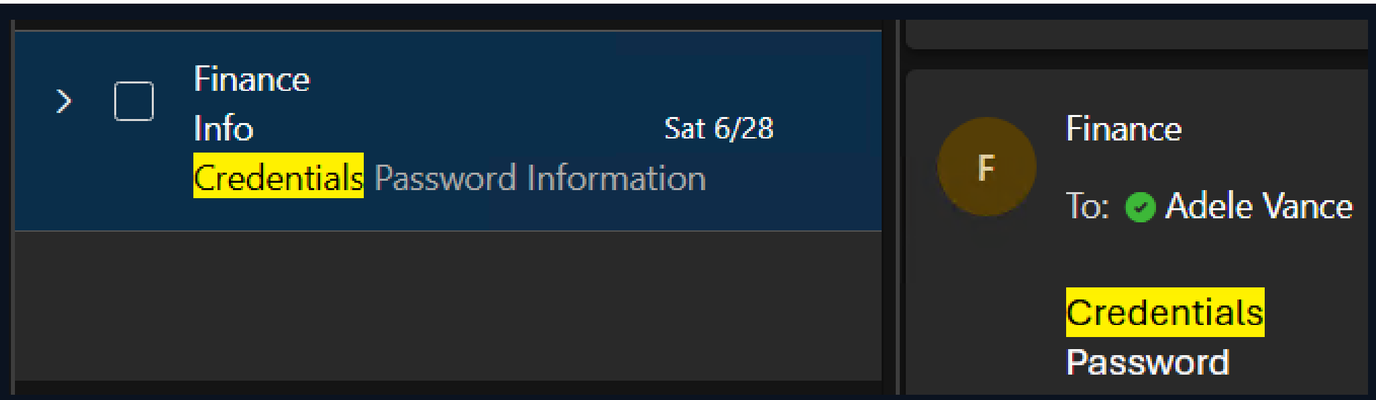
Hidden mail folders

The default value of the `isHidden` property is `false`. You can set **isHidden** only once when creating the mailFolder. You can't update the property using a PATCH operation. To change the **isHidden** property of a folder, delete the existing folder and create a new one with the desired value.

Hidden mail folders support all operations that are supported by a regular mail folder.

Soteria: Hidden Mailbox Folders

Search




```
PS /opt/tools/GraphRunner> Invoke-SearchMailbox -Tokens $tokens -SearchTerm "credentials"
[*] Using the provided access tokens.
[*] Found 9 matches for search term credentials
Subject: Info | Sender: finance@azpurple.com | Receivers: Adele Vance | Date: 06/28/2025 17:59:30
| Message Preview: Credentials Password Information ...
================================================================================
Subject: Dev Environment | Sender: /O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPD
LT)/CN=RECIPIENTS/CN=E286EAE0D96F4FC89B9FD61675ED286F-903988D0-C8 | Receivers: ippsec | Date: 03/0
9/2025 19:31:32 | Message Preview: ... Below are the temporary credentials to the developer enviro
nment. User: mark Pass: MDRisTh3best! Host: 192.168.1.12  Thanks, Adele ...
================================================================================
[*] Do you want to download these emails and their attachments? (Yes/No)
Yes
```

Monitor

- **MailItemsAccessed** (MIA) operation
- Unified Audit Log


```
"AppAccessContext": {
  "APIId": "a3883eba-fbe9-48bd-9ed3-dca3e0e84250",
  "ClientAppId": "a3883eba-fbe9-48bd-9ed3-dca3e0e8
  "IssuedAtTime": "2025-05-13T23:24:55",
  "UniqueTokenId": "a814bde1-22c9-47d9-8e77-42be4c
```

```
"MailboxOwnerUPN": "odie@azpurple.com",
"OperationProperties": [
  {
    "Name": "MailAccessType",
    "Value": "Bind"
  }
],
"OrganizationName": [redacted]
"OriginatingServer": "DM6PR10MB3530 (15.20.4200.000)\r\n",
"TokenTenantId": "afe78b42-40d1-4119-8f59-9a0d3170464c",
"Folders": [
  {
    "FolderItems": [
      {
        "ClientRequestId": "12940c5d-c02f-4e62-a060-53ab1c
        "Id": "RgAAAACtH/ClTkoJTaoW6gkVk5ZQBwCzjxAxm30WR6d
        "ImmutableId": "LgAAAAAdhAMRqmYRzZvIAKoAL8RaDQCzjx
        "InternetMessageId": "<838378c0-46e1-47e3-88ba-7bd
        "SizeInBytes": 164938
      }
    ],
    "Id": "LgAAAACtH/ClTkoJTaoW6gkVk5ZQAQCzjxAxm30WR6dGnjr
    "Path": "\\Inbox"
  }
],
"OperationCount": 1
```

Process

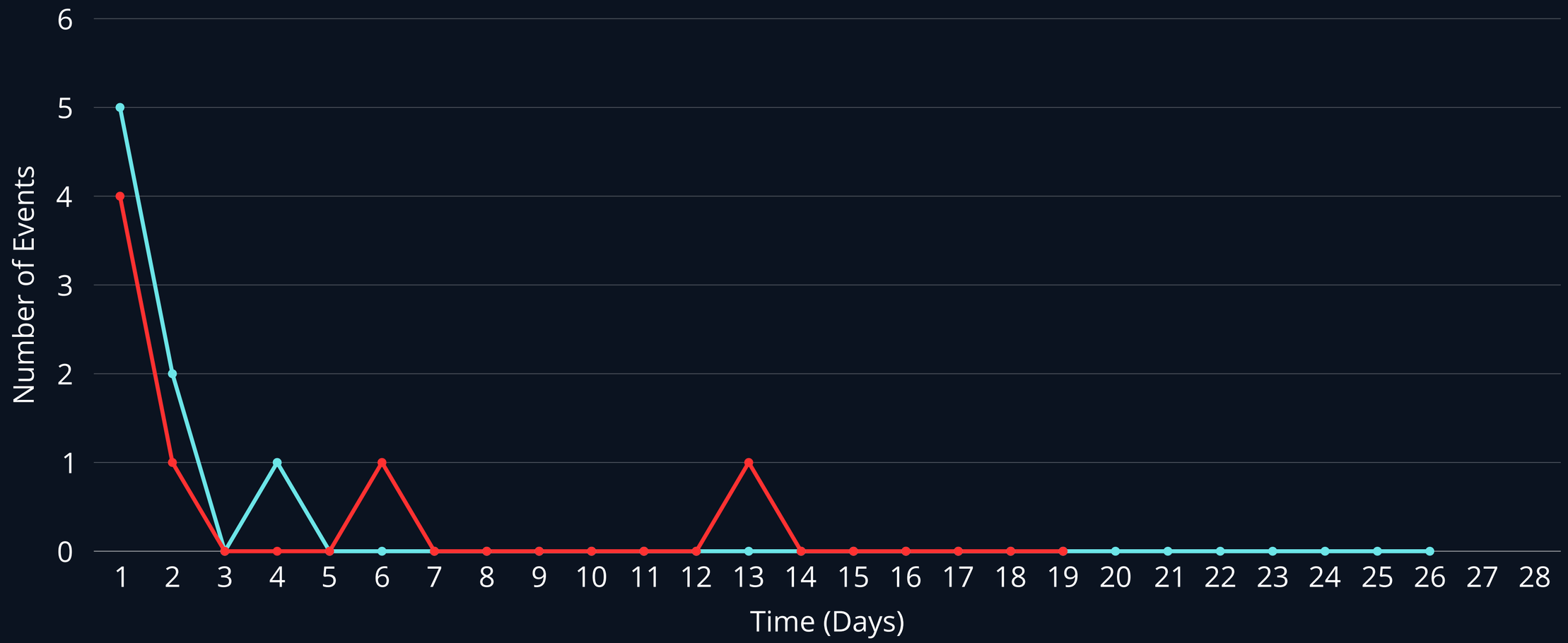
1. Create **hidden folder**
2. Send an email with **keywords**
3. Move email to hidden folder
4. Monitor **MIA** events
5. Alert when email is **accessed**

Results

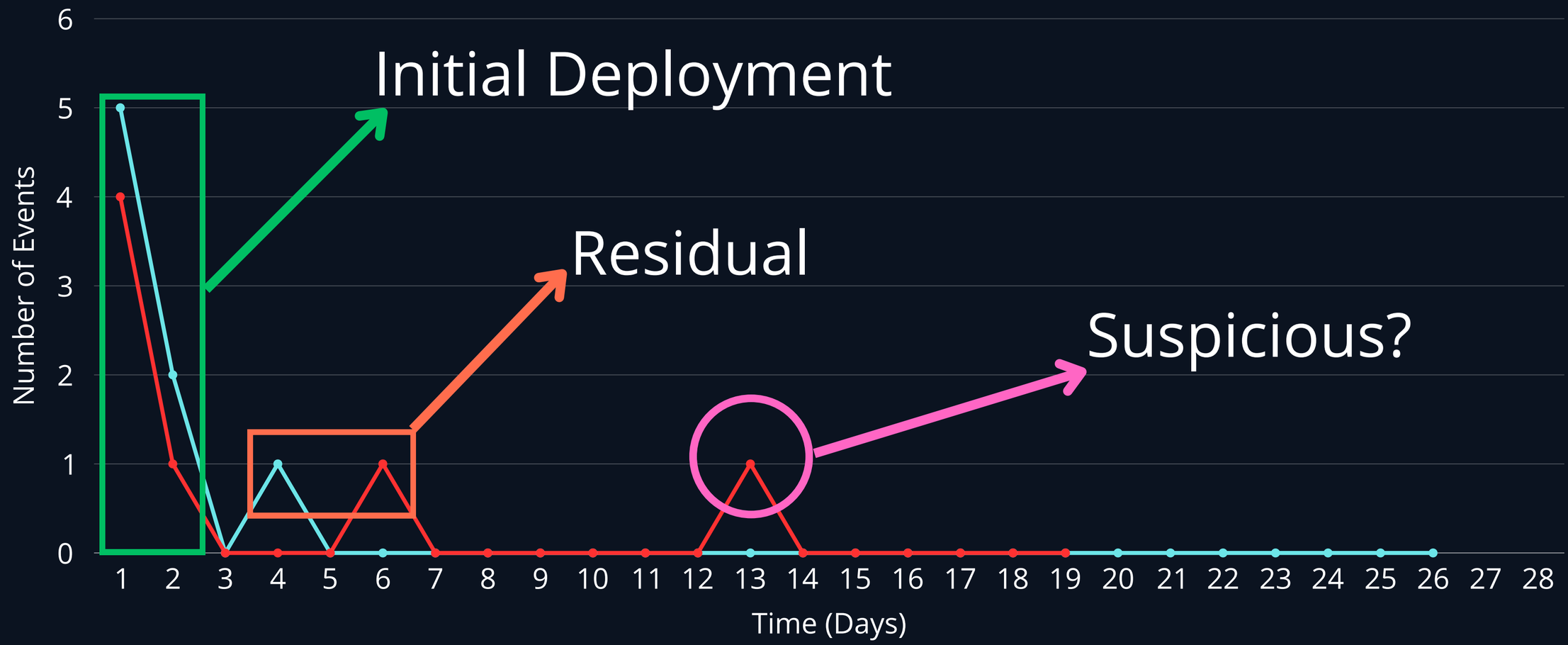
Testing

- Less than 50 employees
- Mix of engineers/sales
- Cloud-native
- **Subset** of users
- 2 different groups
- 4 week trial each

Results


Number of Events
0
1
2
3
4
5
6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Time (Days)

Results


Initial Deployment
Residual
Suspicious?
Number of Events
0
1
2
3
4
5
6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Time (Days)

Conclusion

Criteria - Graded

Attractive → Outlook

Non-Invasive → Hidden Folders

Native → Unified Audit Log

High Fidelity → Almost!

Evolution

And that’s the shift: attackers aren’t hacking computers anymore. They’re hacking trust relationships, identities, and APIs. The whole idea of detection and response needs to evolve with that. Otherwise, we’re securing the hell out of endpoints while attackers happily fish through mailboxes and cloud shares from halfway across the planet.

Florian Roth: @cyb3rops

Questions?

Feel free to reach out!

Contact: linkedin.com/in/odonnell-ryan/

Twitter/X: @odiesec