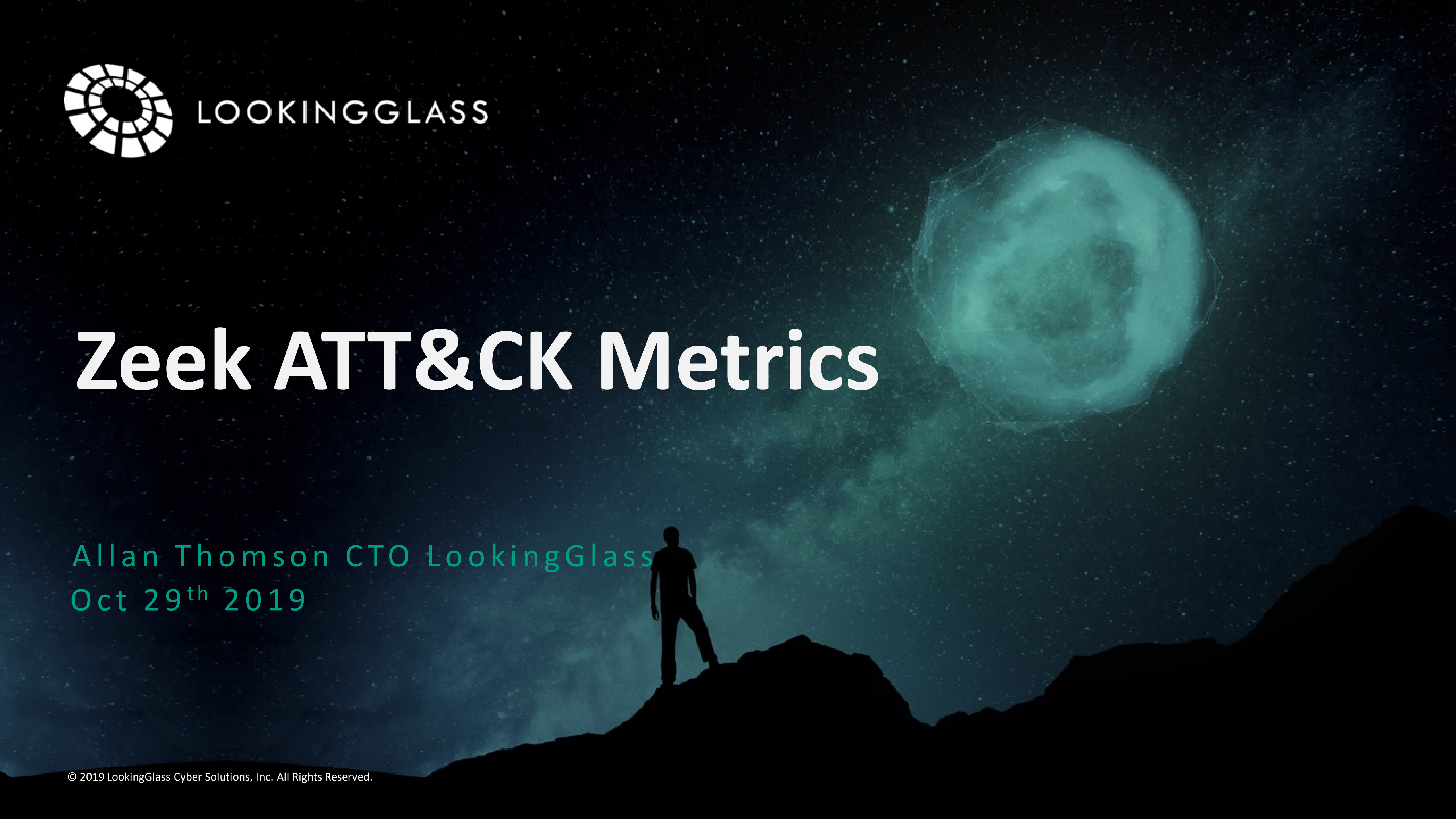
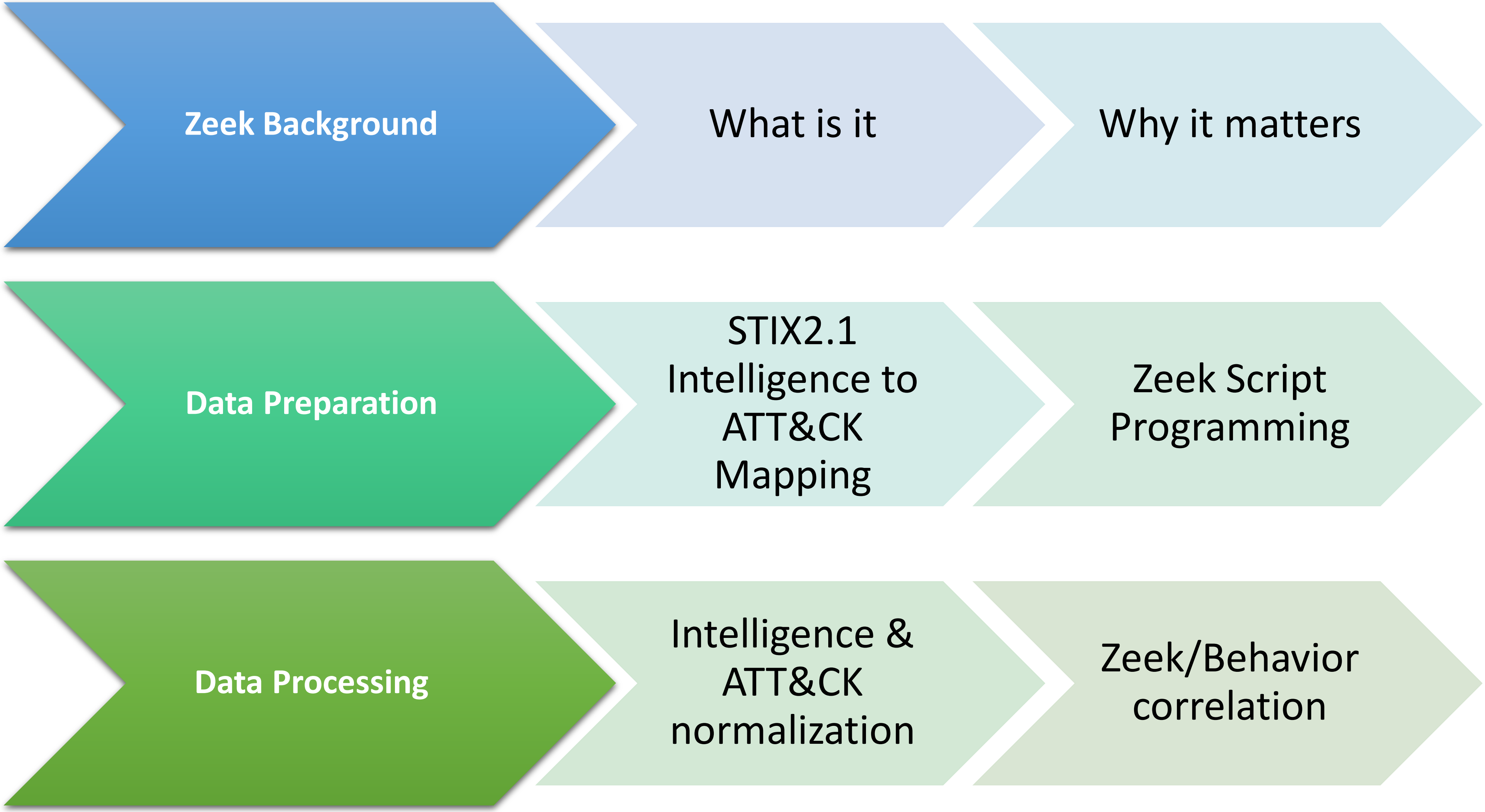


Zeek ATT&CK Metrics

Allan Thomson CTO LookingGlass

Oct 29th 2019

3 Things



Zeek Background
What is it
Why it matters

Data Preparation
STIX2.1 Intelligence to ATT&CK Mapping
Zeek Script Programming

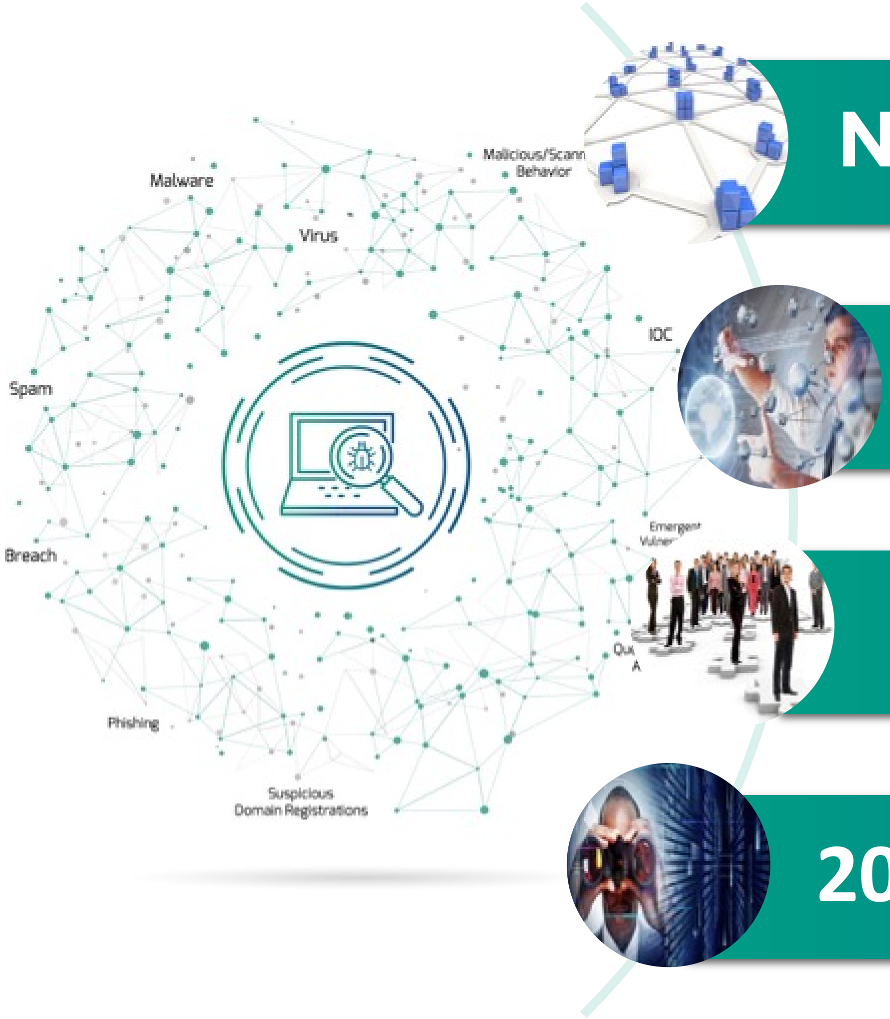
Data Processing
Intelligence & ATT&CK normalization
Zeek/Behavior correlation

LOOKINGGLASS

Zeek Background

Basic Architecture

Zeek Background



Network Analysis Framework

Focused on **Network Security Monitoring**

Open Source Community

20 Years Research (www.zeek.org)

Zeek Ecosystem

Connections

RPC

NTLM

Spam

ICMP

Bittorrent

IRC

Botnet

Syslog

Exfiltration

Protocol Vulnerabilities

Geo-location

APT

X509 Certificates

Statistics

Bruteforce

Web - HTTP

Sandbox Integration

Payload

Email

Scanning

File Sharing - SMB

SSH

IPv4

NTP

Certificate Validation

Bitcoin

DHCP

Shellshock

VirusTotal Integration

Routing - RIP

Intelligence Integration

Domain - DNS

SSL/TLS

MAC

Fast Flux

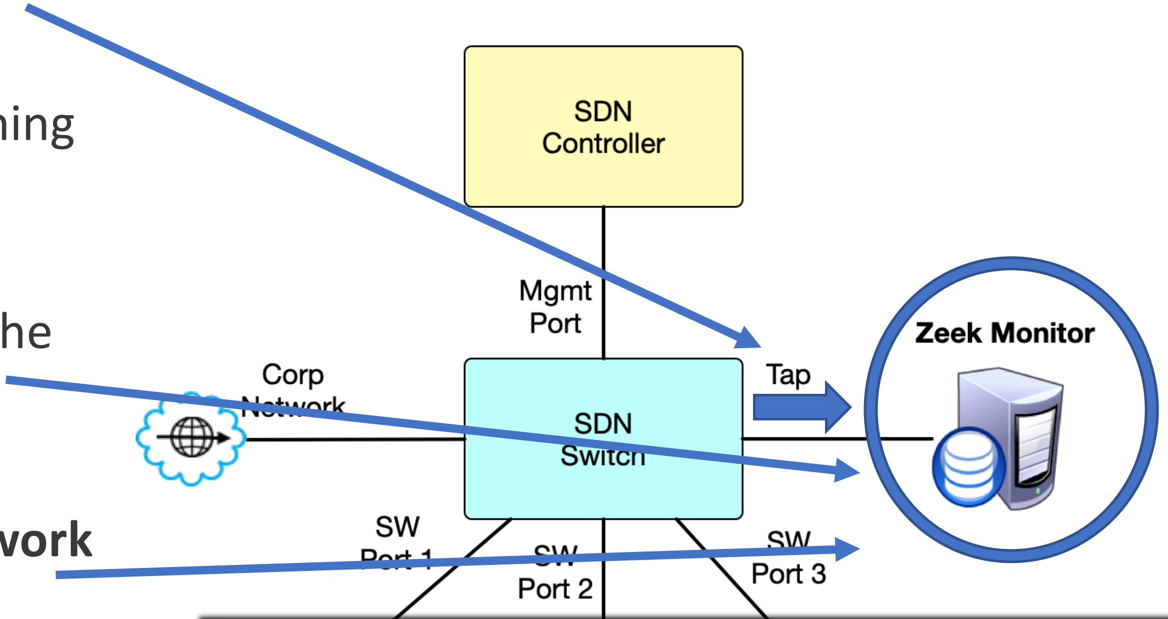
Blacklists

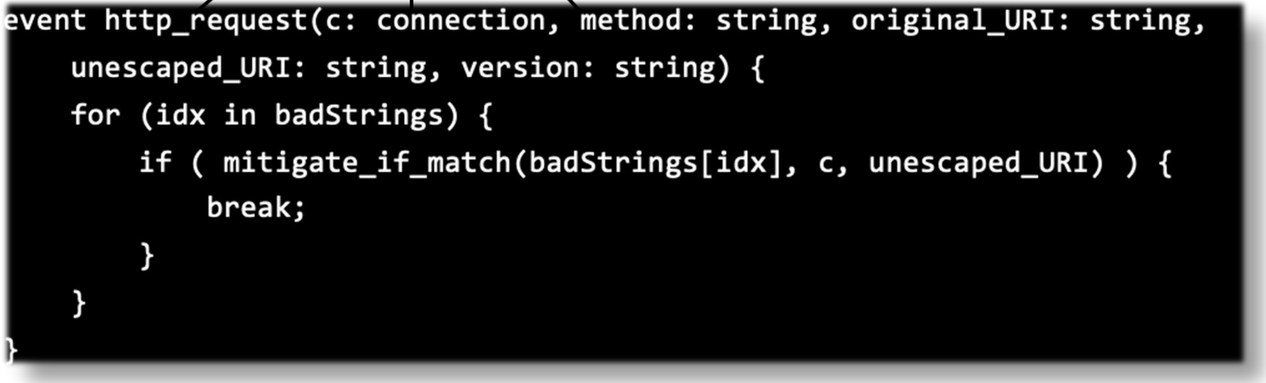
IPv6

Wordpress

Zeek Based Detection

- Zeek monitor receives copy of all traffic
- Zeek employs an event-based programming model
- Zeek scripts run to perform analysis on the network traffic
- Identify stateful analysis on specific **network patterns** or **network behavior**
- Can also identify **user application behaviors (i.e. nefarious activity)**



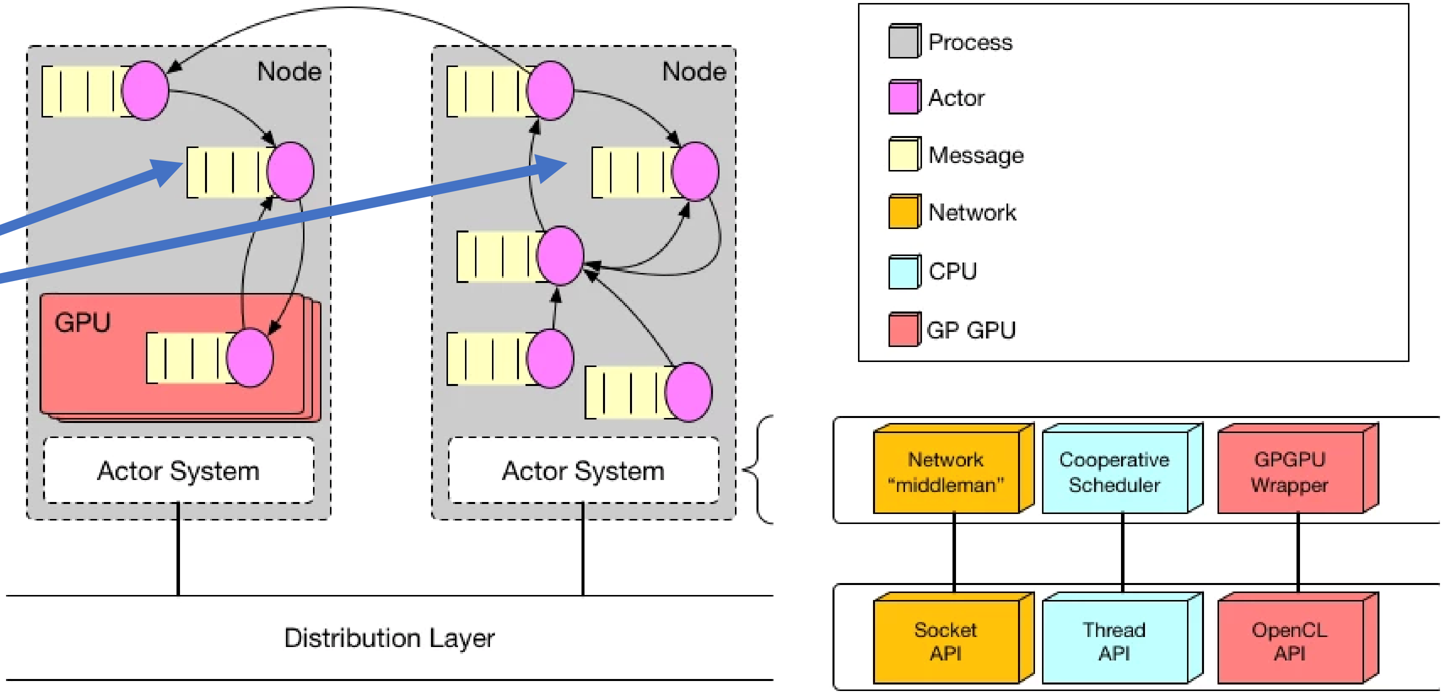


```
event http_request(c: connection, method: string, original_URI: string,
    unescaped_URI: string, version: string) {
    for (idx in badStrings) {
        if ( mitigate_if_match(badStrings[idx], c, unescaped_URI) ) {
            break;
        }
    }
}
```

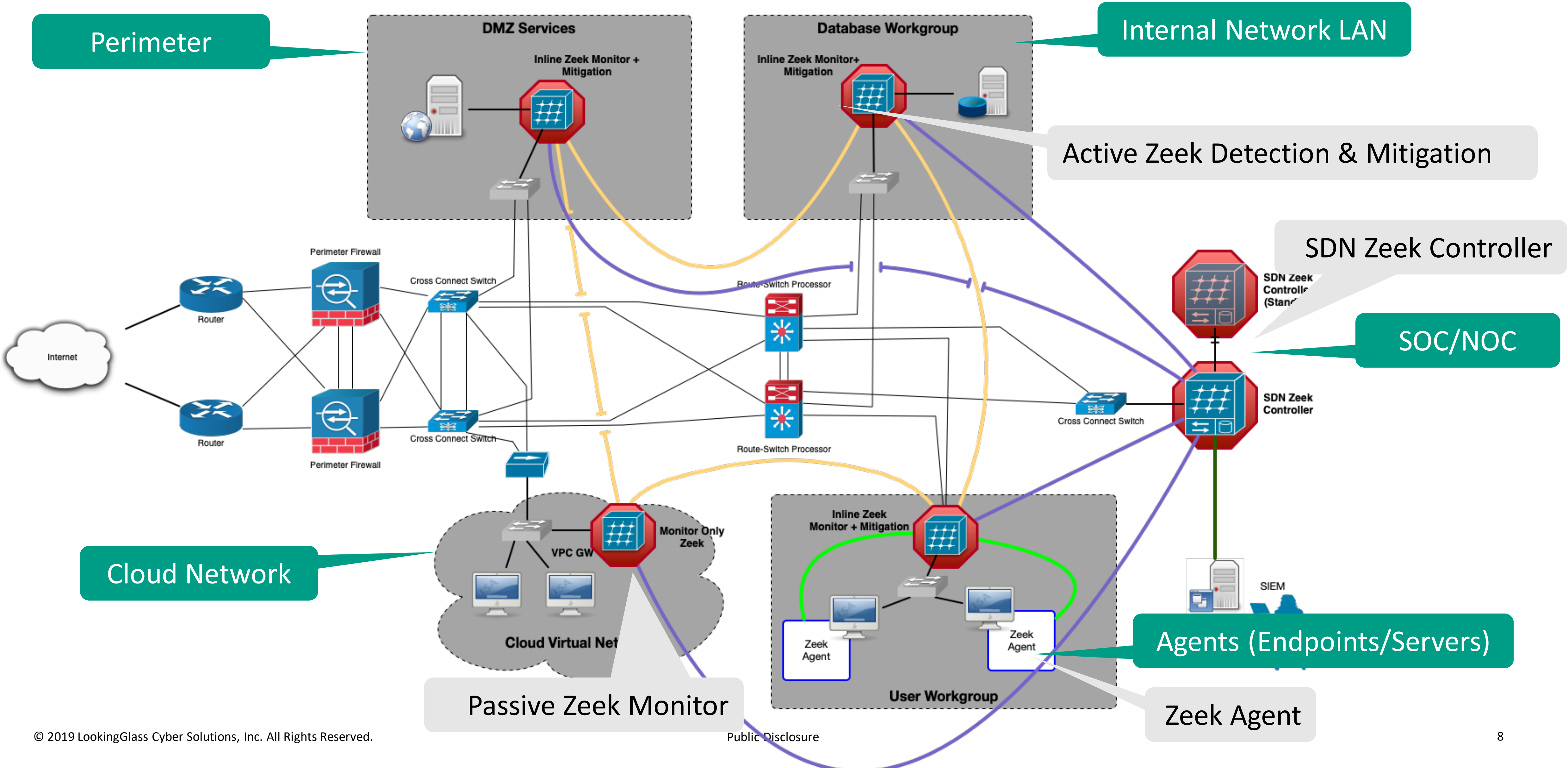
Zeek Processing and Distribution

- Supports Actor-Framework (https://actor-framework.org/)

- Distributed **Messaging** & **Processing**
  - Event Processing
  - Cross-Event Correlation
  - Behavioral Identification
  - Intelligence Correlation
  - **ATT&CK Analysis**
    - **Multi-node**
    - **Multi-processing**



Applying Zeek For ATT&CK/Intelligence Correlation



Perimeter

Internal Network LAN

Active Zeek Detection & Mitigation

SDN Zeek Controller

SOC/NOC

Cloud Network

Agents (Endpoints/Servers)

Passive Zeek Monitor

Zeek Agent

LOOKINGGLASS

Data Preparation

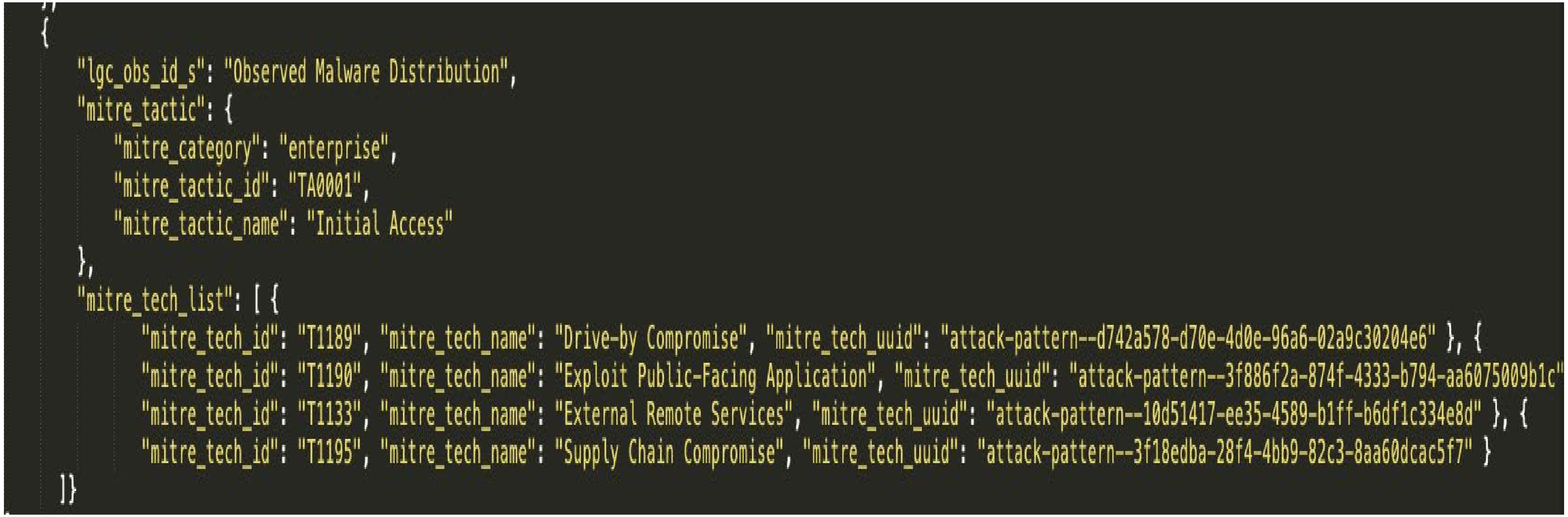
Intelligence & Zeek Updates for ATT&CK

Threat Intelligence & ATT&CK

- How we modelled Threat Intelligence → STIX2
- How we related Intelligence to ATT&CK → STIX2
- How we correlated intel with activities (net, sys, user) → Zeek
- How we applied action based on Intelligence/ATT&CK → Zeek

Data Preparation: Intel to ATT&CK Mapping

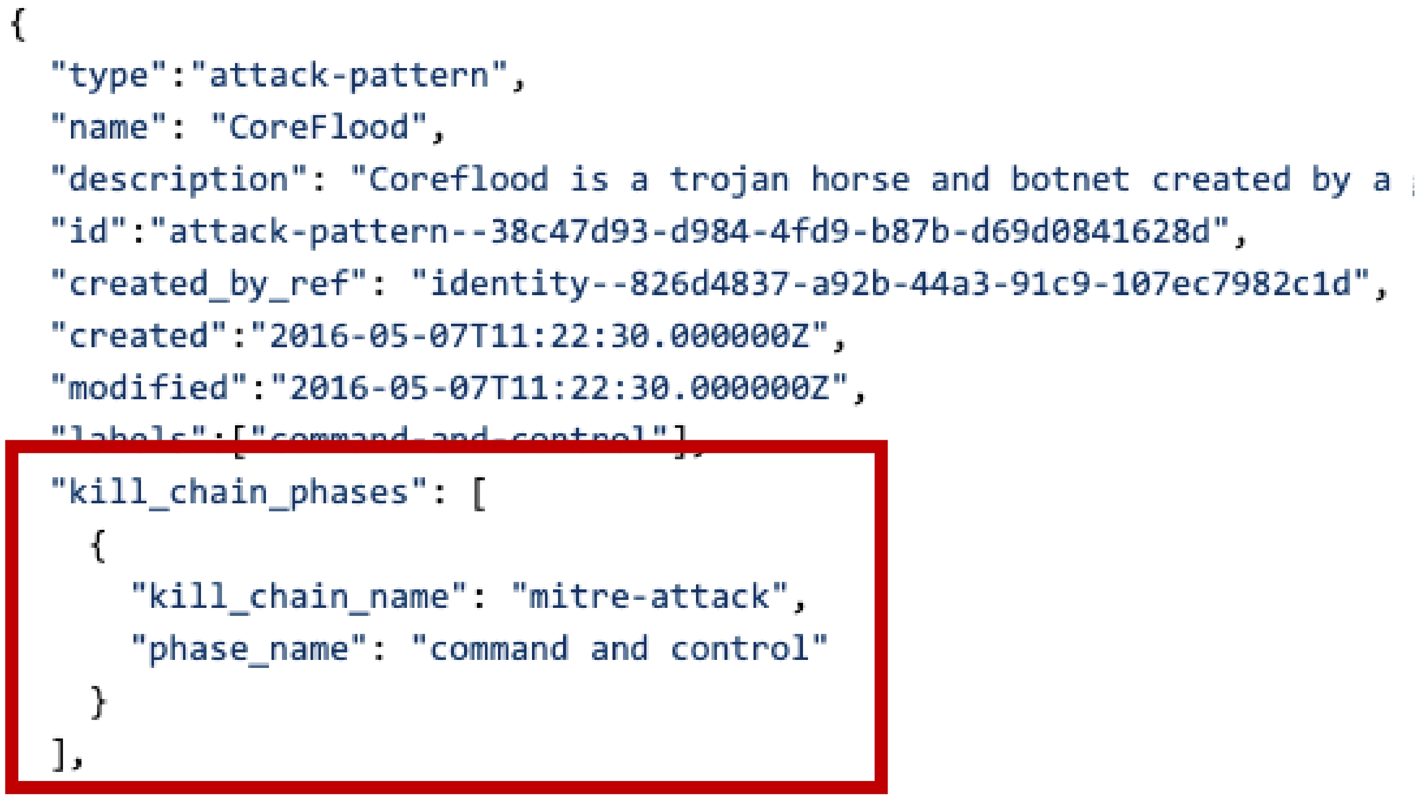
- 90 different intelligence feeds
- ~1800 ***Unique*** intelligence attack-patterns, intrusion sets, actors
  - Data-driven Mapping to ATT&CK
  - Include ATT&CK Mapping when producing STIX2.1 Intelligence



```
  {
    "lgc_obs_id_s": "Observed Malware Distribution",
    "mitre_tactic": {
      "mitre_category": "enterprise",
      "mitre_tactic_id": "TA0001",
      "mitre_tactic_name": "Initial Access"
    },
    "mitre_tech_list": [ {
        "mitre_tech_id": "T1189", "mitre_tech_name": "Drive-by Compromise", "mitre_tech_uuid": "attack-pattern--d742a578-d70e-4d0e-96a6-02a9c30204e6" }, {
        "mitre_tech_id": "T1190", "mitre_tech_name": "Exploit Public-Facing Application", "mitre_tech_uuid": "attack-pattern--3f886f2a-874f-4333-b794-aa6075009b1c"
        "mitre_tech_id": "T1133", "mitre_tech_name": "External Remote Services", "mitre_tech_uuid": "attack-pattern--10d51417-ee35-4589-b1ff-b6df1c334e8d" }, {
        "mitre_tech_id": "T1195", "mitre_tech_name": "Supply Chain Compromise", "mitre_tech_uuid": "attack-pattern--3f18edba-28f4-4bb9-82c3-8aa60dcac5f7" }
   ]}
```

Data Preparation: Intel to ATT&CK Mapping

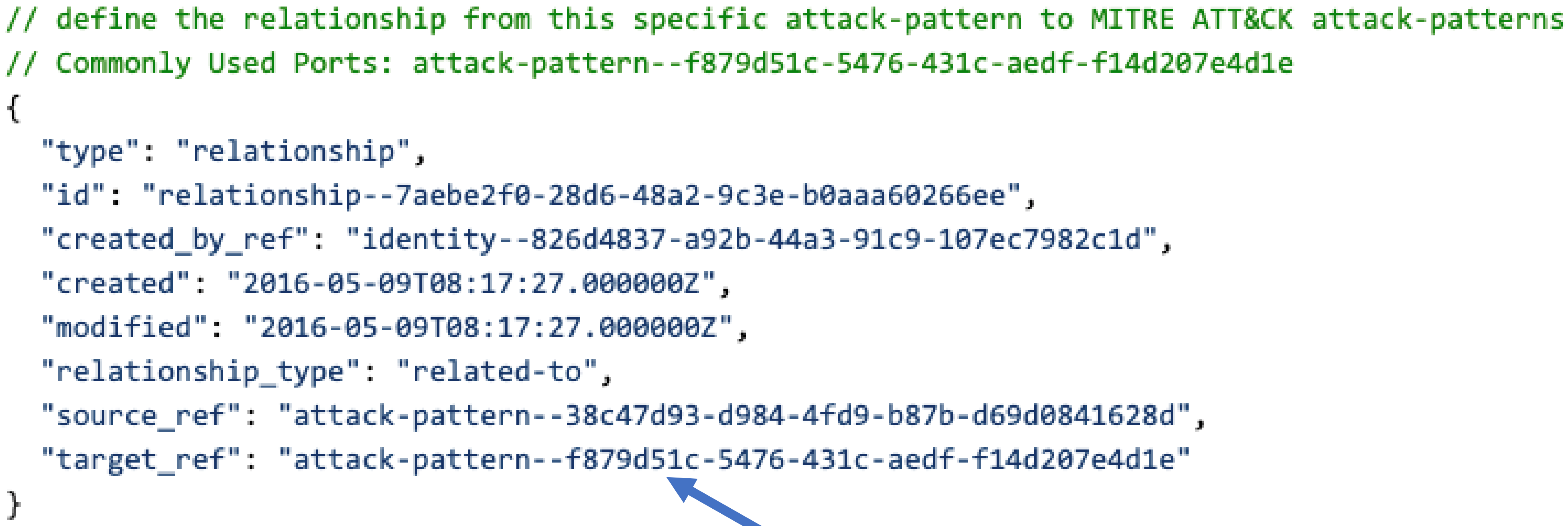
- **Tactics** mapped using kill-chain property on Intel Feed
  - Attack-Pattern SDO
  - Intrusion Set SDO
  - Actor SDO



```
{
  "type":"attack-pattern",
  "name": "CoreFlood",
  "description": "Coreflood is a trojan horse and botnet created by a
  "id":"attack-pattern--38c47d93-d984-4fd9-b87b-d69d0841628d",
  "created_by_ref": "identity--826d4837-a92b-44a3-91c9-107ec7982c1d",
  "created":"2016-05-07T11:22:30.000000Z",
  "modified":"2016-05-07T11:22:30.000000Z",
  "labels":["command-and-control"],
  "kill_chain_phases": [
    {
      "kill_chain_name": "mitre-attack",
      "phase_name": "command and control"
    }
  ],
```

Data Preparation: Intel to ATT&CK Mapping

- Intel Feed Attack-Patterns related to ATT&CK **Attack-Patterns** using SROs



```
// define the relationship from this specific attack-pattern to MITRE ATT&CK attack-patterns
// Commonly Used Ports: attack-pattern--f879d51c-5476-431c-aedf-f14d207e4d1e
{
  "type": "relationship",
  "id": "relationship--7aebe2f0-28d6-48a2-9c3e-b0aaa60266ee",
  "created_by_ref": "identity--826d4837-a92b-44a3-91c9-107ec7982c1d",
  "created": "2016-05-09T08:17:27.000000Z",
  "modified": "2016-05-09T08:17:27.000000Z",
  "relationship_type": "related-to",
  "source_ref": "attack-pattern--38c47d93-d984-4fd9-b87b-d69d0841628d",
  "target_ref": "attack-pattern--f879d51c-5476-431c-aedf-f14d207e4d1e"
}
```

Points to ATT&CK UUID

Data Processing

Correlation, Alerting

An Intelligence Question

- Find ***IP ranges*** and ***CIDRs*** that are associated with the ***NAICS Industry of 'Carpet and Rug Mills';***
  - discover **all active IPs** contained within these ranges,
  - and
    - find ***FQDNs*** associated with them where those FQDNs have ***active threats***
    - *that include*
      - ***Attack-Pattern Exploitation of Remotes Services*** *and*
      - ***Attack-Pattern Pass the Hash***

The problem answering that question?

- Many different sources assert **essentially the same data**
  - i.e. *FeedA asserts that IP 10.0.0.1 has Malware A, and FeedB asserts the same*

- Much of the metadata is the **same across temporal series**
  - Repeated fact assertions and threat associations
  - i.e. *FeedB asserts that Actor BB, associated with Intrusion Set AA, using Attack-Pattern ZZZ Drive-by Compromise malware YY on Infrastructure CC at time X, and again, at time Y*

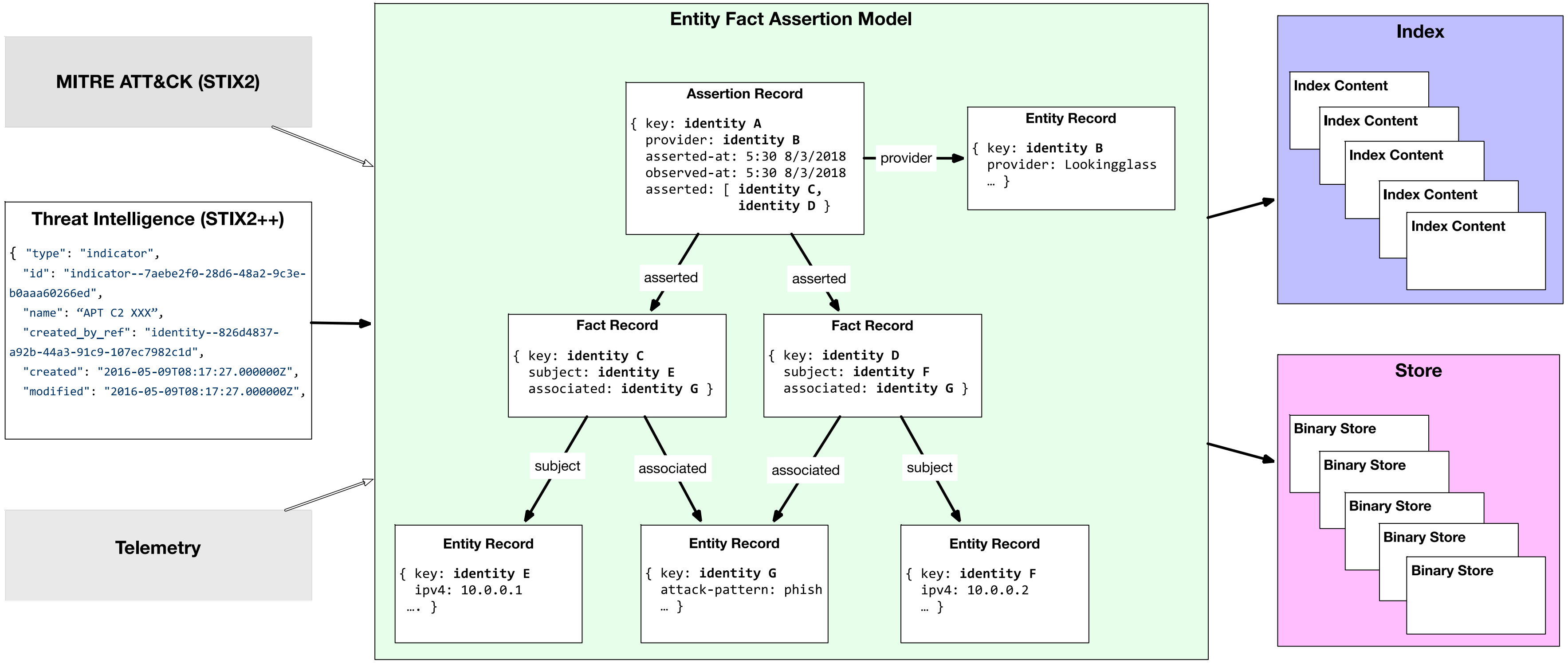
- **Different attributes** with **different data representation** that communicate the **same semantic information**
  - *i.e. country_s of "United States" and "United States of America" and country_code_s of "US", and "USA"*

- Multiple different object/entity types, ***billions*** of instances that requires ***large-scale join*** across data-sets where those data-sets are being updated in ***real-time***

Solution: Unified Data Modelling

- **Entity**
  - Contains information about an **Entity** that will never change
  - Metadata such as name, and IP ranges in **Entities** allows **Facts** to only contain reference

- **Fact**
  - Asserts attributes on **Entity** and relationships to other **Entities**
  - If **Facts** contained temporal and source/provider attributes, it would be multiple **Fact Record** for each



Threat Intelligence (STIX2++)

```
{ "type": "indicator",
  "id": "indicator--7aebe2f0-28d6-48a2-9c3e-
b0aaa60266ed",
  "name": “APT C2 XXX”,
  "created_by_ref": "identity--826d4837-
a92b-44a3-91c9-107ec7982c1d",
  "created": "2016-05-09T08:17:27.000000Z",
  "modified": "2016-05-09T08:17:27.000000Z",
```

Assertion Record

```
{ key: identity A
  provider: identity B
  asserted-at: 5:30 8/3/2018
  observed-at: 5:30 8/3/2018
  asserted: [ identity C,
              identity D }
```

Entity Record

```
{ key: identity B
  provider: Lookingglass
  … }
```

Fact Record

```
{ key: identity C
  subject: identity E
  associated: identity G }
```

Fact Record

```
{ key: identity D
  subject: identity F
  associated: identity G }
```

Entity Record

```
{ key: identity E
  ipv4: 10.0.0.1
  …. }
```

Entity Record

```
{ key: identity G
  attack-pattern: phish
  … }
```

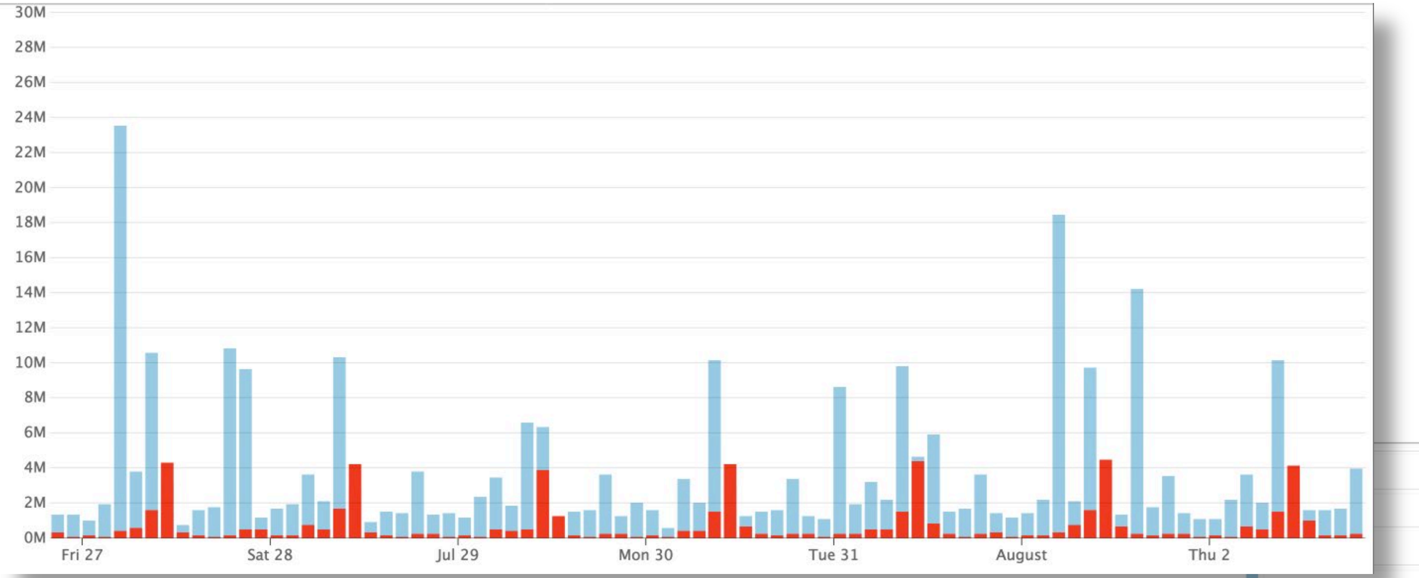
Entity Record

```
{ key: identity F
  ipv4: 10.0.0.2
  … }
```

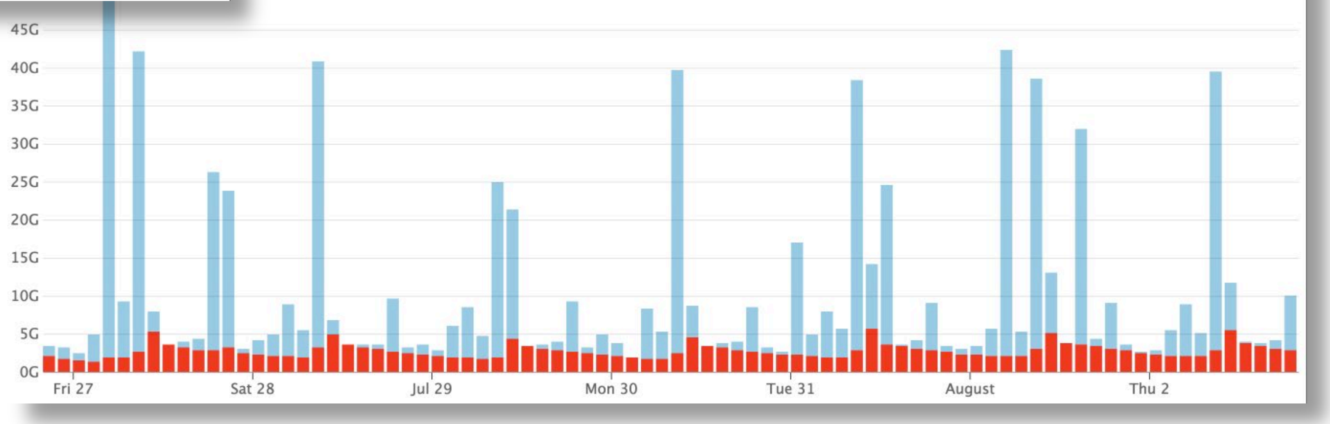
- **Assertion**
  - Asserts one or more **Facts** by **Source Entity** and **Provider Entity**
  - Contains all temporal attributes – **Observed At,** and **Asserted At**

Impact on Data

Records: ~150mm/day → ~19mm/day.

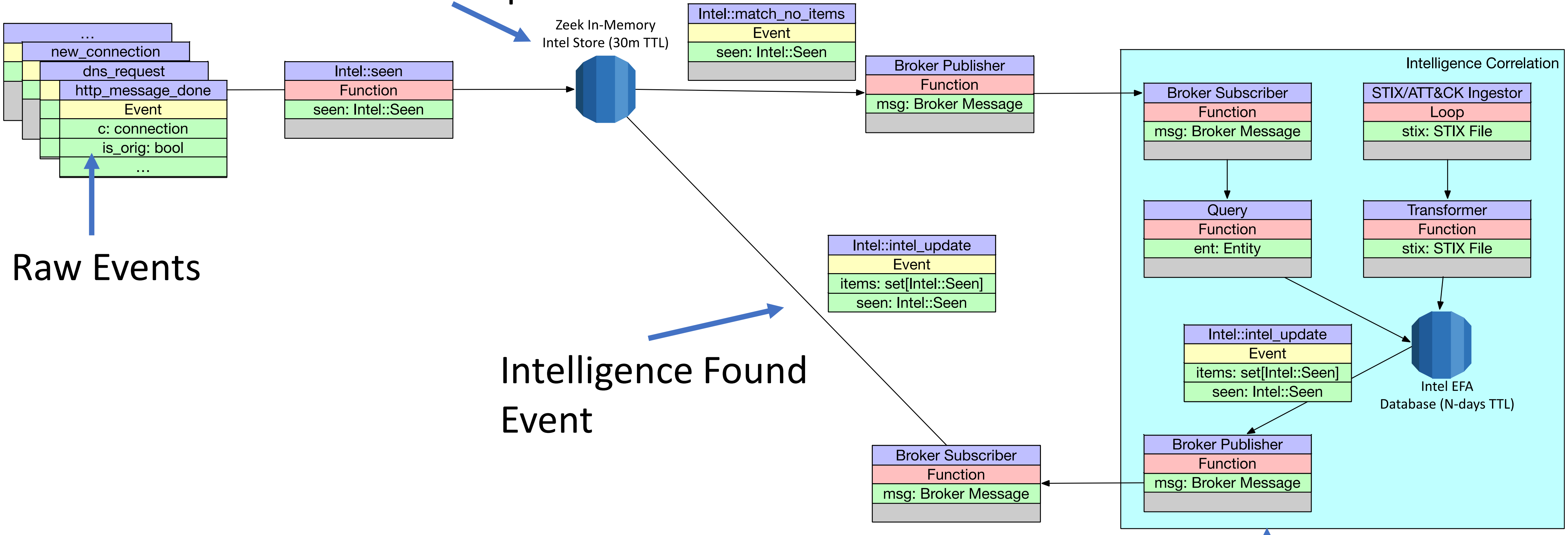


Bytes: ~150GB/day → ~25-30GB/day.



Zeek Intelligence Basic Lookup Pipeline

Zeek Intelligence Framework Lookup



...

new_connection

dns_request

http_message_done
Event
c: connection
is_orig: bool
...

Intel::seen
Function
seen: Intel::Seen

Zeek In-Memory
Intel Store (30m TTL)

Intel::match_no_items
Event
seen: Intel::Seen

Broker Publisher
Function
msg: Broker Message

Intel::intel_update
Event
items: set[Intel::Seen]
seen: Intel::Seen

Broker Subscriber
Function
msg: Broker Message

Intelligence Correlation

Broker Subscriber
Function
msg: Broker Message

STIX/ATT&CK Ingestor
Loop
stix: STIX File

Query
Function
ent: Entity

Transformer
Function
stix: STIX File

Intel::intel_update
Event
items: set[Intel::Seen]
seen: Intel::Seen

Intel EFA
Database (N-days TTL)

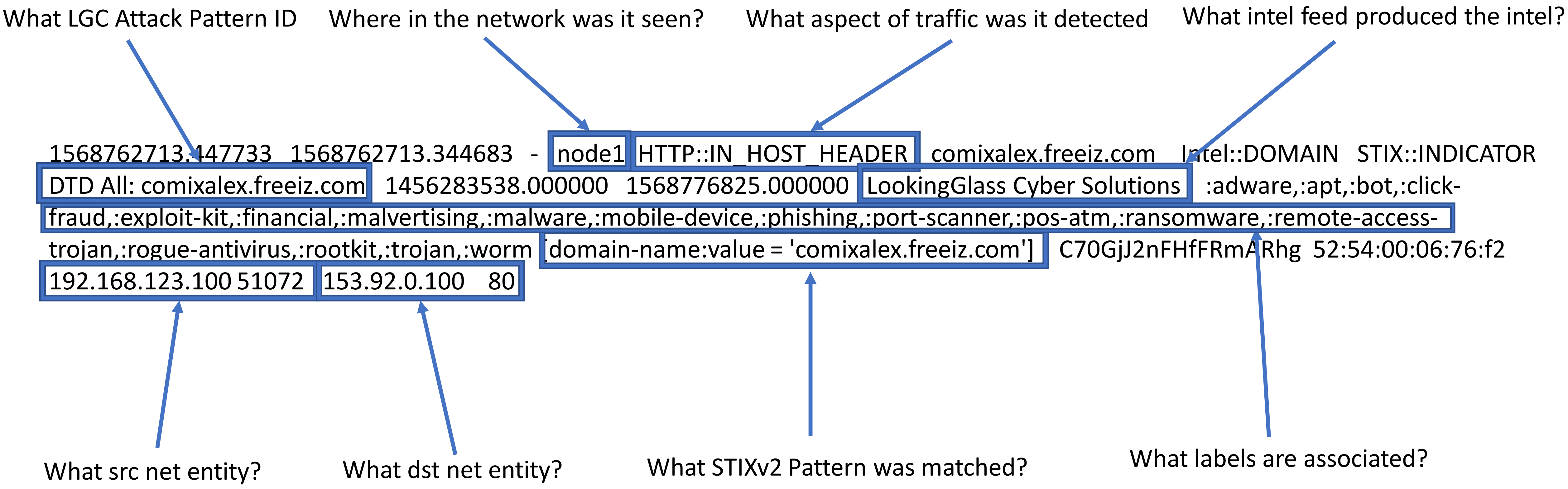
Broker Publisher
Function
msg: Broker Message

Raw Events

Intelligence Found Event

Intelligence & ATT&CK Correlation

Zeek ATT&CK Report Event Dissection



What LGC Attack Pattern ID

Where in the network was it seen?

What aspect of traffic was it detected

What intel feed produced the intel?

```
1568762713.447733   1568762713.344683   -  node1  HTTP::IN_HOST_HEADER   comixalex.freeiz.com    Intel::DOMAIN   STIX::INDICATOR
DTD All: comixalex.freeiz.com   1456283538.000000   1568776825.000000  LookingGlass Cyber Solutions   :adware,:apt,:bot,:click-
fraud,:exploit-kit,:financial,:malvertising,:malware,:mobile-device,:phishing,:port-scanner,:pos-atm,:ransomware,:remote-access-
trojan,:rogue-antivirus,:rootkit,:trojan,:worm  [domain-name:value = 'comixalex.freeiz.com']   C70GjJ2nFHfFRmARhg  52:54:00:06:76:f2
192.168.123.100 51072  153.92.0.100    80
```

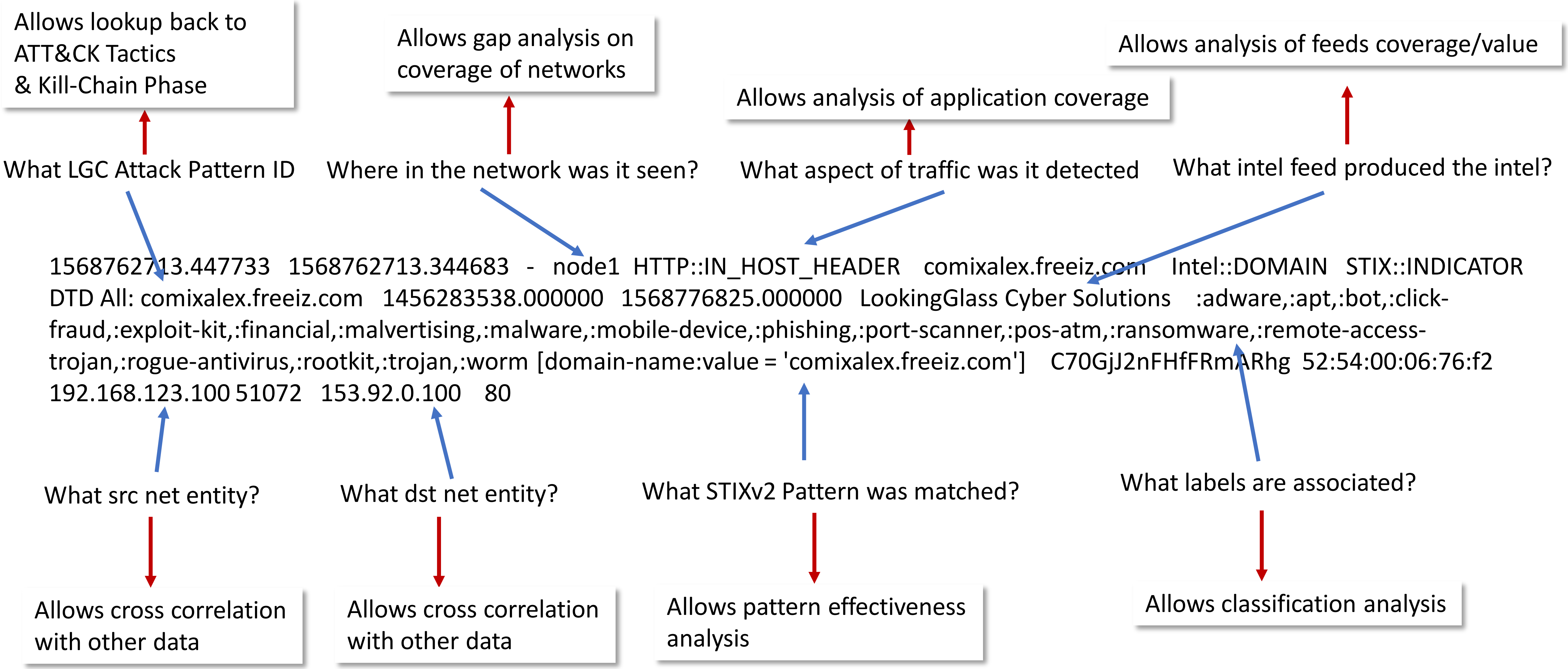
What src net entity?

What dst net entity?

What STIXv2 Pattern was matched?

What labels are associated?

Zeek ATT&CK Report Event Analysis



Allows lookup back to ATT&CK Tactics & Kill-Chain Phase

Allows gap analysis on coverage of networks

Allows analysis of application coverage

Allows analysis of feeds coverage/value

What LGC Attack Pattern ID

Where in the network was it seen?

What aspect of traffic was it detected

What intel feed produced the intel?

```
1568762713.447733   1568762713.344683   -   node1  HTTP::IN_HOST_HEADER    comixalex.freeiz.com    Intel::DOMAIN   STIX::INDICATOR
DTD All: comixalex.freeiz.com   1456283538.000000   1568776825.000000   LookingGlass Cyber Solutions    :adware,:apt,:bot,:click-
fraud,:exploit-kit,:financial,:malvertising,:malware,:mobile-device,:phishing,:port-scanner,:pos-atm,:ransomware,:remote-access-
trojan,:rogue-antivirus,:rootkit,:trojan,:worm [domain-name:value = 'comixalex.freeiz.com']    C70GjJ2nFHfFRmARhg  52:54:00:06:76:f2
192.168.123.100 51072   153.92.0.100    80
```

What src net entity?

What dst net entity?

What STIXv2 Pattern was matched?

What labels are associated?

Allows cross correlation with other data

Allows cross correlation with other data

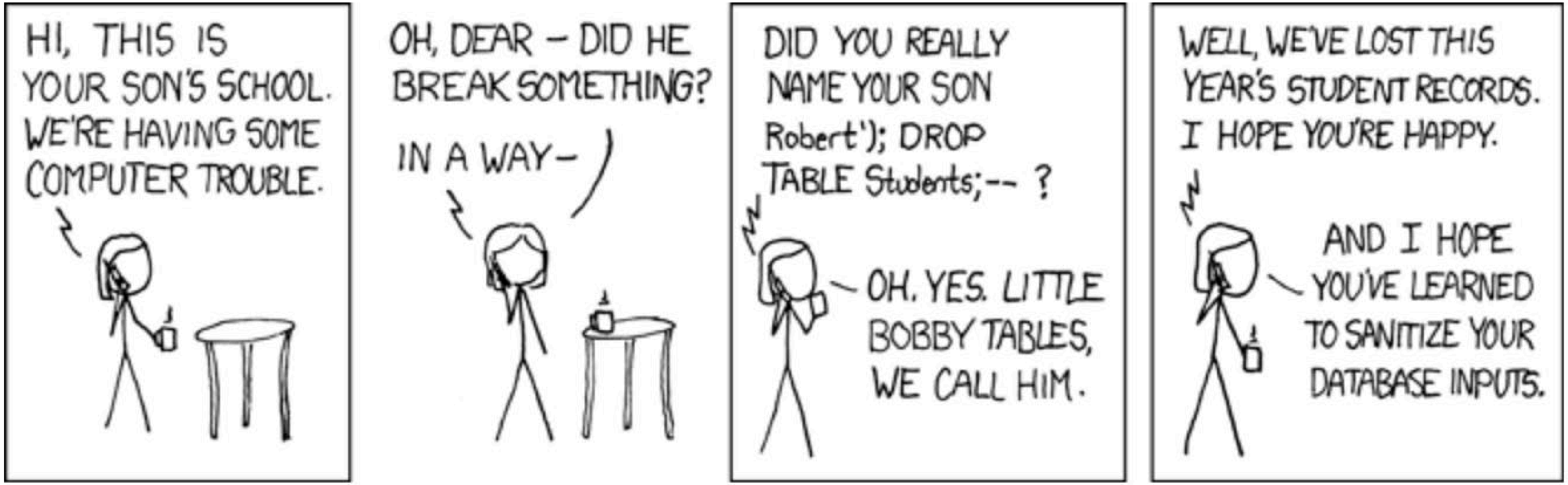
Allows pattern effectiveness analysis

Allows classification analysis

Summary

- Zeek provides **effective** and **flexible** framework for **collection** and **correlation**
- Data **preparation** & **modelling** can have big **impact** on analysis effectiveness
- Data **correlation** at scale requires **end-to-end** approach

Questions?


HI, THIS IS YOUR SON'S SCHOOL. WE'RE HAVING SOME COMPUTER TROUBLE.
OH, DEAR – DID HE BREAK SOMETHING?
IN A WAY–
DID YOU REALLY NAME YOUR SON Robert'); DROP TABLE Students;-- ?
OH. YES. LITTLE BOBBY TABLES, WE CALL HIM.
WELL, WE'VE LOST THIS YEAR'S STUDENT RECORDS. I HOPE YOU'RE HAPPY.
AND I HOPE YOU'VE LEARNED TO SANITIZE YOUR DATABASE INPUTS.

LOOKINGGLASS

Thank You

LookingGlassCyber.com

@LookingGlassCyber

@LG_Cyber

/company/LookingGlass

YouTube

/LookingGlassCyber