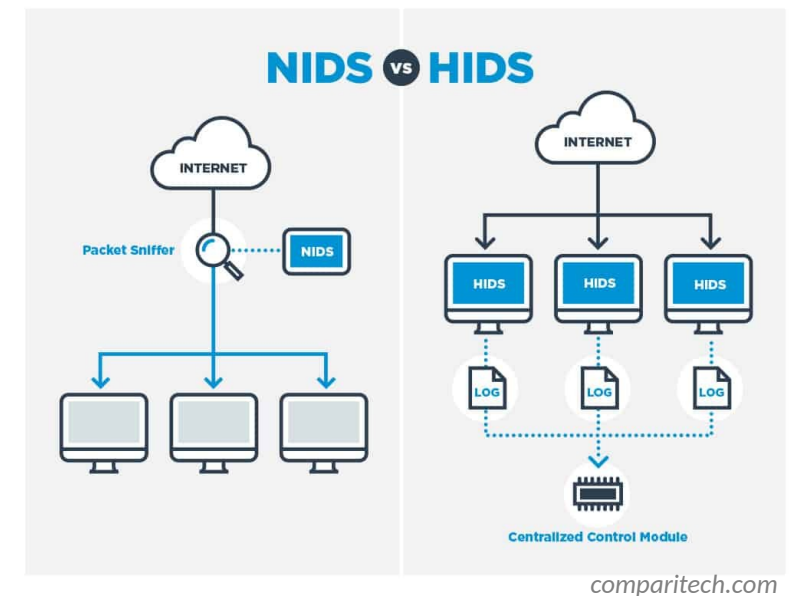


Network Traffic Analysis with Malcolm

Seth Grover, Malcolm developer • Cybersecurity R&D • Idaho National Lab

Intrusion Detection Systems

- HIDS: Host Intrusion Detection Systems
 - Agents run on individual hosts or devices on a network
 - Not what we're talking about today
- NIDS: Network Intrusion Detection Systems
 - Monitor and analyze network traffic for anomalies: suspicious activity, policy violations, etc.
 - Generally passive/out-of-band; otherwise it's an Intrusion Prevention System
 - Detection methods
 - Signature-based detection (e.g., Suricata)
 - Statistical anomaly-based detection (e.g., Random Cut Forest)
 - Stateful protocol analysis detection (e.g., Zeek)





IDS: Types of Attacks

- Scanning Attack
 - Determine network topology
 - IDS highlights connections from one host to many other hosts in the network, or connection attempts to sequential IP addresses and/or ports
- Denial of Service Attack
 - Interrupt service by flooding requests or flaws in protocol implementations
 - IDS identifies large volume of traffic from or to a particular host or invalid connection states (e.g., TCP SYN/ACK with no ACK)
- Penetration Attack
 - Gain access to system resources by exploiting a software or configuration flaw
 - Trickier, but IDS may detect vulnerable software versions or simply alert on unusual operations (e.g., a “write” operation in an already-configured environment with mostly “read” operations)





- Extensible, open-source passive network analysis framework
- More than just an Intrusion Detection System:
 - Packet capture (like **TCPDUMP**)
 - Traffic inspection (like  Wireshark)
 - Intrusion detection (like **SNORT**)
 - Log recording (like NetFlow and syslog)
 - Scripting framework (like  python™)



Strengths

- Analyzes both link-layer and application-layer behavior
- Content extraction
- Behavioral analysis
- Session correlation
- Can add support for uncommon protocols through scripts/plugins

Weaknesses

- Session metadata only (not full payload)
- Setup and configuration can be complicated
- Produces flat textual log files which can be unwieldy for in-depth analysis

- Network Protocols
- Files
- Detection
- Network Observations

[illegible][illegible]

Network Protocols

- `conn` – Network session tracking
 - Identified by session 4-tuple (originating IP:port, responding IP:port)
 - One session (line in a log file) for every IP connection
 - Unique identifier (UID) ties lines from other logs to a session
- `http`, `modbus`, `ftp`, `dns`, **etc.**
 - Protocol-specific log files created as traffic is seen
 - Contain application-layer metadata about network activities

Files

- `files` – File analysis results
 - Each transferred file identified with FUID
 - Associated with connection UID(s) over which file was transferred
 - File name, mime type, file size, etc. provided when available
- `pe` – Analysis of Portable Executable (PE) files
 - Target platform, architecture, OS, etc. for executables transferred across the network
- `x509` – Analysis of X.509 public key certificates

Detection

- `notice` – Zeek concept of “alarms,” notices draw extra attention to an event
 - `Conn::Content_Gap`, `DNS::External_Name`,
`FTP::Bruteforcing`, `Heartbleed::SSL_Heartbeat_Attack`,
`HTTP::SQL_Injection_Attacker`, `Scan::Address_Scan`,
`Scan::Port_Scan`, `Software::Vulnerable_Version`,
`SSH::Password_Guessing`, `SSL::Certificate_Expired`,
`Weird::Activity`, ...
 - <https://docs.zeek.org/en/stable/zeek-noticeindex.html>

Detection (cont.)

- `weird` – Unexpected network-level activity
 - > 150 weirdness indicators across many protocols
 - <https://docs.zeeb.org/en/stable/scripts/base/frameworks/notice/weird.zeeb.html#id1>
- `signatures` – Signature matches, including hits from enabled carved file scanners like ClamAV, YARA and capa

Network Observations

- Periodic dump of entities seen over the last day
 - `known_certs` – SSL certificates
 - `known_devices` – MAC addresses
 - `known_hosts` – Hosts with TCP handshakes
 - `known_modbus` – Modbus masters and slaves
 - `known_services` – Services (TCP “servers”)
 - `software` – Software being used on the network (e.g., Apache, OpenSSH, etc.)
 - Could be used for identifying vulnerable versions of software or firmware



Arkime

Strengths

- Large scale index packet capture and search tool
- Packet analysis engine with support for many common IT protocols
- Web interface for browsing, searching, analysis and PCAP carving for exporting
- PCAP payloads (not just session header/metadata) are viewable and searchable

Weaknesses

- No OT protocol support
- Adding new protocol parsers requires C programming

Malcolm

A powerful open-source network traffic analysis tool suite.

<https://github.com/idaholab/Malcolm>



Streamlined deployment

- Suitable for field use (hunt or incident response) or SOC deployment. Runs in Docker on Linux, macOS and Windows platforms. Provides easy-to-use web-based user interfaces.

Industry-standard tools

- Uses Arkime and Zeek for network traffic capture, Logstash for parsing and enrichment, OpenSearch for indexing and Dashboards and Arkime Viewer for visualization. Also leverages OpenSearch Anomaly Detection, Suricata IDS, YARA, capa, ClamAV, CyberChef and other proven tools for analysis of traffic and artifacts.

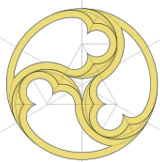
Expanding control systems visibility

- Analyzes more protocols used in operational technology (OT) networks than other open-source or paid solutions. Ongoing development is focused on increasing the quantity and quality of industrial control systems (ICS) traffic.

Dedicated sensor appliance

- Includes Hedgehog Linux, a hardened Linux distribution for capturing network traffic and forwarding its metadata to Malcolm.

Malcolm



Components

<https://github.com/idaholab/Malcolm/#Components>



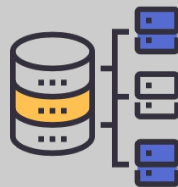
Capture &
Analysis



File Scanning



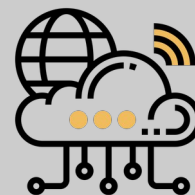
Forwarding &
Enrichment



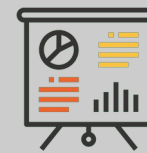
Storage



Anomaly
Detection



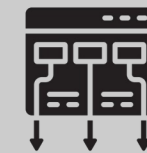
Asset
Management



Visualization



Payload
Analysis



Framework



Anomaly
Detection
Plugin



OpenSearch
Dashboards



CyberChef



OpenSearch



Alerting



Arkime

Arkime
session PCAP
export to

WIRESHARK



beats



Alerting
Plugin



Supported Protocols

<https://github.com/idaholab/Malcolm/#Protocols>

Internet layer

Border Gateway Protocol (BGP)

Building Automation and Control (BACnet)

Bristol Standard Asynchronous Protocol (BSAP)

Distributed Computing Environment / Remote Procedure Calls
(DCE/RPC)

Dynamic Host Configuration Protocol (DHCP)

Distributed Network Protocol 3 (DNP3)

Domain Name System (DNS)

EtherCAT

EtherNet/IP / Common Industrial Protocol (CIP)

FTP (File Transfer Protocol)

GENISYS

Google Quick UDP Internet Connections (gQUIC)

Hypertext Transfer Protocol (HTTP)

IPsec

Internet Relay Chat (IRC)

Lightweight Directory Access Protocol (LDAP)

Kerberos

Modbus

MQ Telemetry Transport (MQTT)

MySQL

NT Lan Manager (NTLM)

Network Time Protocol (NTP)

Oracle

**Open Platform Communications Unified Architecture
(OPC UA) Binary**

Open Shortest Path First (OSPF)

OpenVPN

PostgreSQL

Process Field Net (PROFINET)

Remote Authentication Dial-In User Service (RADIUS)

Remote Desktop Protocol (RDP)

Remote Framebuffer / Virtual Network Computing (RFB/VNC)

S7comm / Connection Oriented Transport Protocol (COTP)

Secure Shell (SSH)

Secure Sockets Layer (SSL) / Transport Layer Security (TLS)

Session Initiation Protocol (SIP)

Server Message Block (SMB) / Common Internet File System (CIFS)

Simple Mail Transfer Protocol (SMTP)

Simple Network Management Protocol (SNMP)

SOCKS

STUN (Session Traversal Utilities for NAT)

Syslog

Tabular Data Stream (TDS)

Telnet / remote shell (rsh) / remote login (rlogin)

TFTP (Trivial File Transfer Protocol)

WireGuard

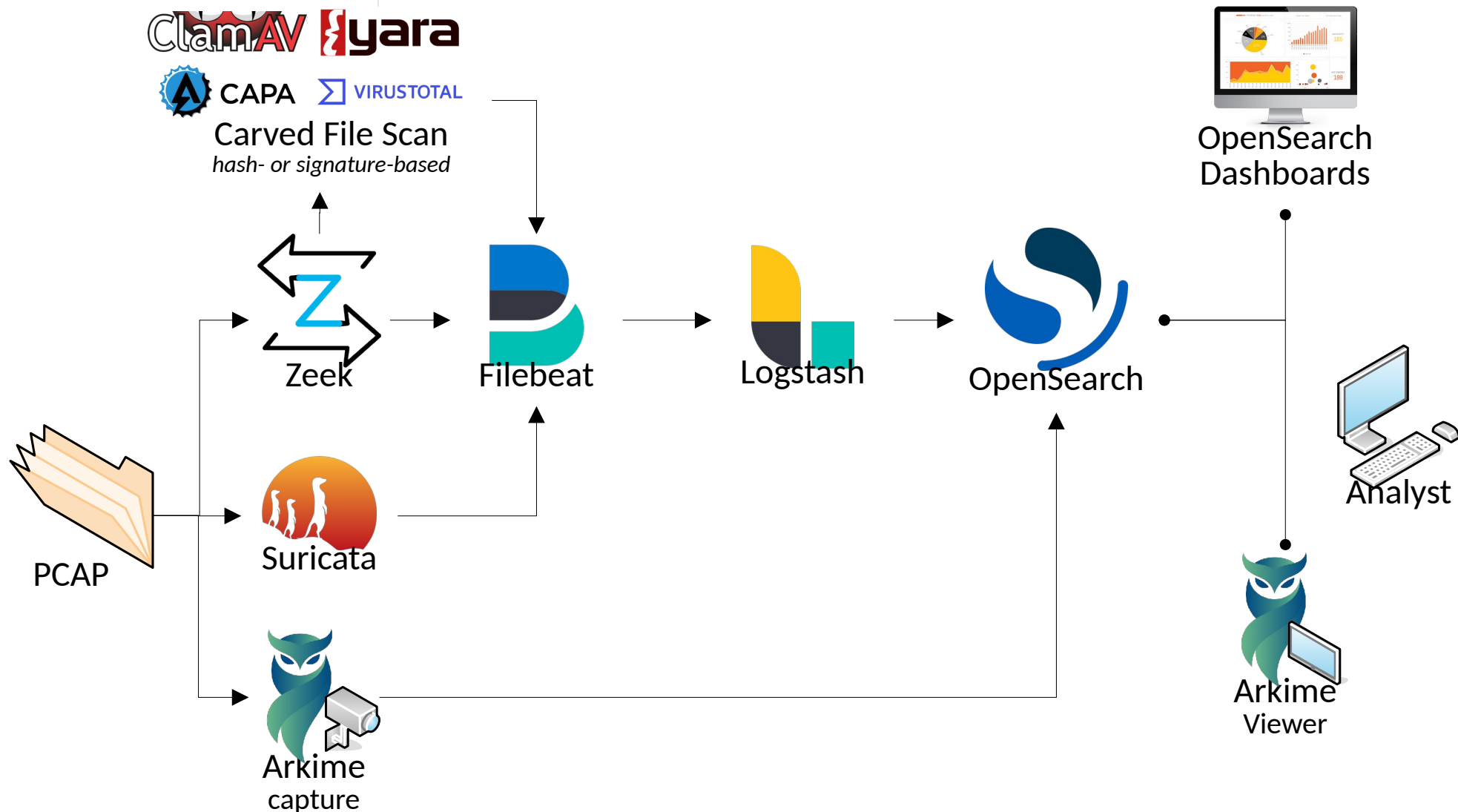
various tunnel protocols (e.g., GTP, GRE, Teredo, AYIYA, IP-in-IP, etc.)

** Industrial control systems protocols indicated with **bold***

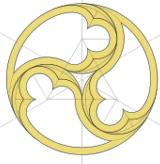
Malcolm

Data Pipeline

<https://github.com/idaholab/Malcolm>



Malcolm



Data Pipeline

<https://github.com/idaholab/Malcolm>

Traffic is collected passively by the Hedgehog sensor device

- Zeek, Arkime Capture and Suricata generate metadata about network communications
- Full PCAP may be stored locally on the sensor
- Files transfers are detected and the files scanned for threats
- PCAP may also be uploaded to or captured by Malcolm without requiring a dedicated sensor

Metadata is securely forwarded to Malcolm

- All communications between the sensor and aggregator are TLS-encrypted
- Sensor data including resource utilization, syslog, audit logs, temperatures and more may also be forwarded

Logs are enriched and stored in OpenSearch

- Lookups are performed for GeoIP, ASN, MAC-to-vendor, community ID, domain name entropy, etc.
- Network events normalized across protocols and data sources
- Best-guess techniques applied for identifying obscure ICS traffic
- Enriched metadata may be forwarded to higher-tiered Malcolm instance

Machine learning algorithms identify anomalies

- Default detectors are provided for action and result, flow size and types of transferred files
- Custom detectors may be created for any aspect of any supported protocol

Alerts are sent over email, webhooks, Slack or Amazon Chime

- Alerts may be triggered by exceeded thresholds, anomalies detected, custom queries, etc.

Traffic is visualized in OpenSearch Dashboards and Arkime Viewer

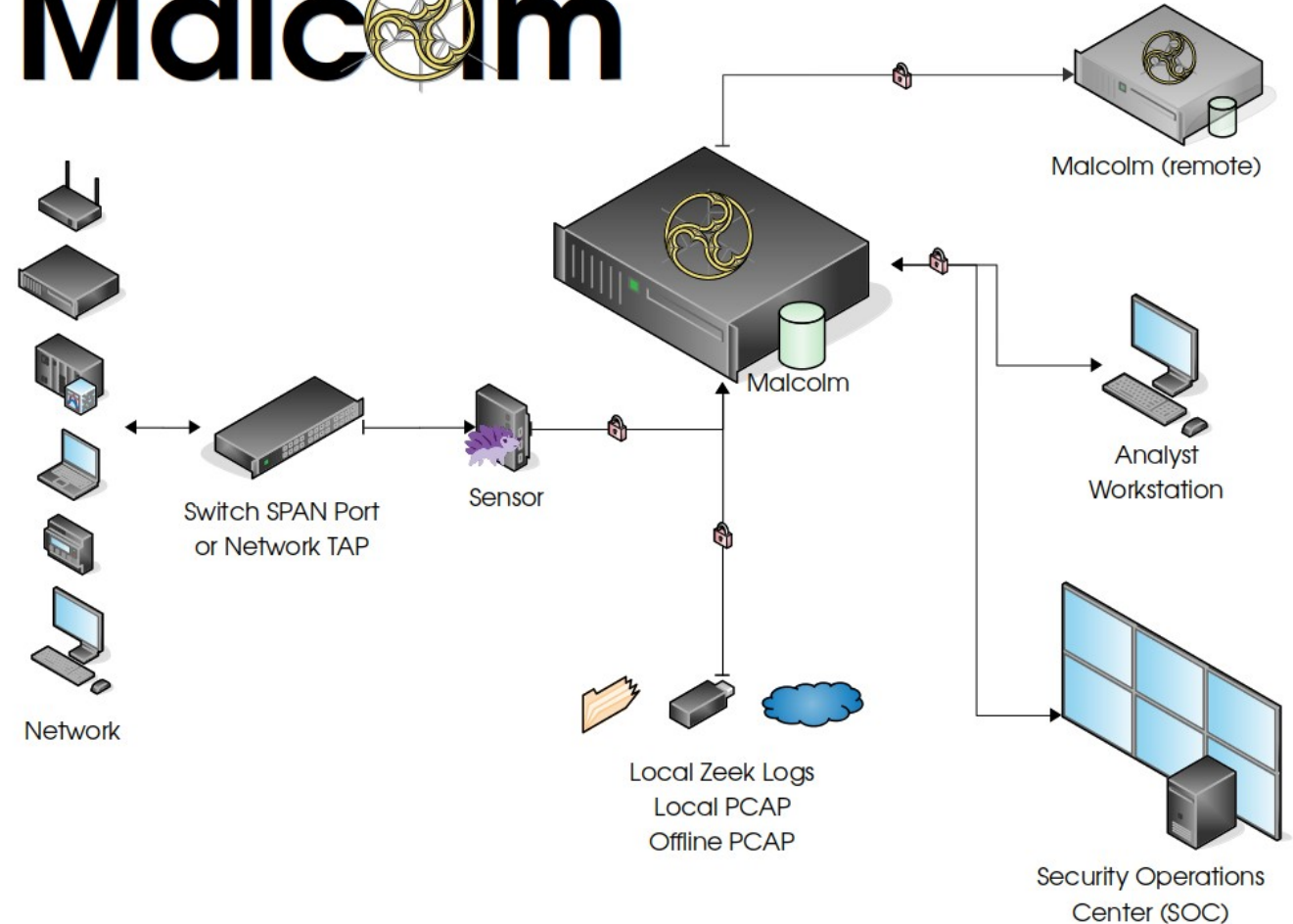
- Dozens of custom dashboards are provided for all supported protocols
- PCAP payloads are retrieved from sensor automatically on demand
- Custom visualizations may be created via drag-and-drop interface
- Malcolm can authenticate users from its own list or via Active Directory / LDAP

Configuring and Running Malcolm

- Runs natively in Docker or in a Virtual Machine
- 16+GB RAM, 4+ cores, “enough” disk for PCAP and logs suggested
- Documentation and source code on GitHub: github.com/idaholab/Malcolm
- Walkthroughs on [YouTube](#): search “Malcolm Network Traffic Analysis”



Malcolm









Identifying Network Hosts and Subnets

- Assign custom names to network hosts and subnets prior to PCAP import
- Allows identification of cross-segment traffic and name-based search and filter
- Define in text file(s) or via web interface
- <https://localhost/name-map-ui>



The screenshot shows a web interface for mapping network addresses to names. It features a table with columns for Address, Name, and Tag. Each row includes edit and delete icons. A search bar is at the top right, and a save button is at the bottom right.

	Address	Name	Tag		
	06:46:0b:a6:16:bf	serial-host.intranet.lan	testbed		
ent	10.0.0.0/8	corporate			
	127.0.0.1	localhost			
	127.0.1.1	localhost			
ent	172.16.0.0/12	virtualized	testbed		
	192.168.10.10	office-laptop.intranet.lan			
ent	192.168.40.0/24	corporate			
ent	192.168.50.0/24	corporate			
ent	192.168.100.0/24	control			
ent	192.168.200.0/24	dmz			
	:::1	localhost			

▼

Address

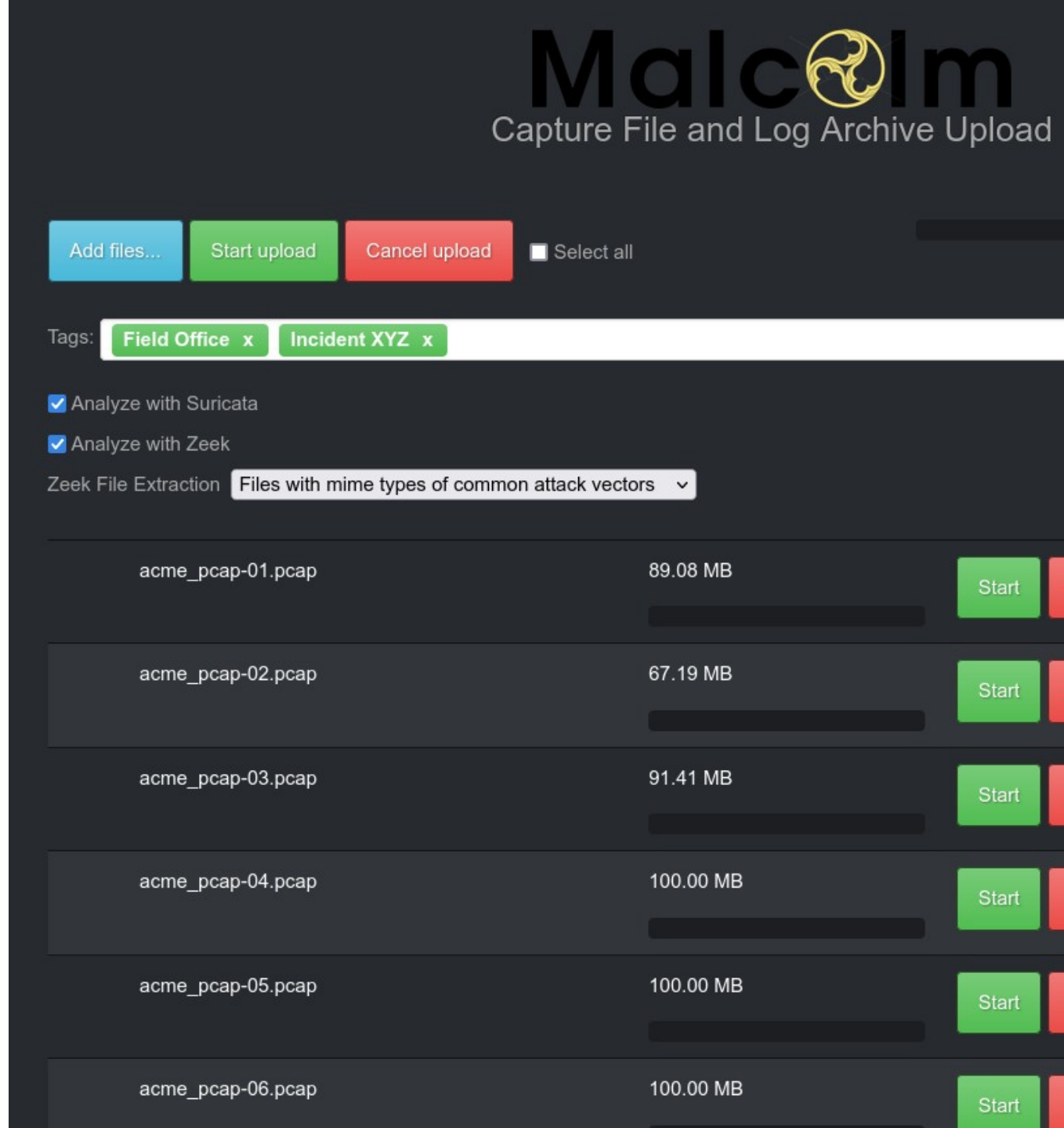
Name

Tag (optional)



Importing Traffic Captures for Analysis

- Specify tags for search and filter
- Enable Suricata and/or Zeek analysis and file extraction
 - Or configure as global defaults
- Upload PCAP files or archived Zeek logs
 - pcapng not supported yet
- <https://localhost/upload>



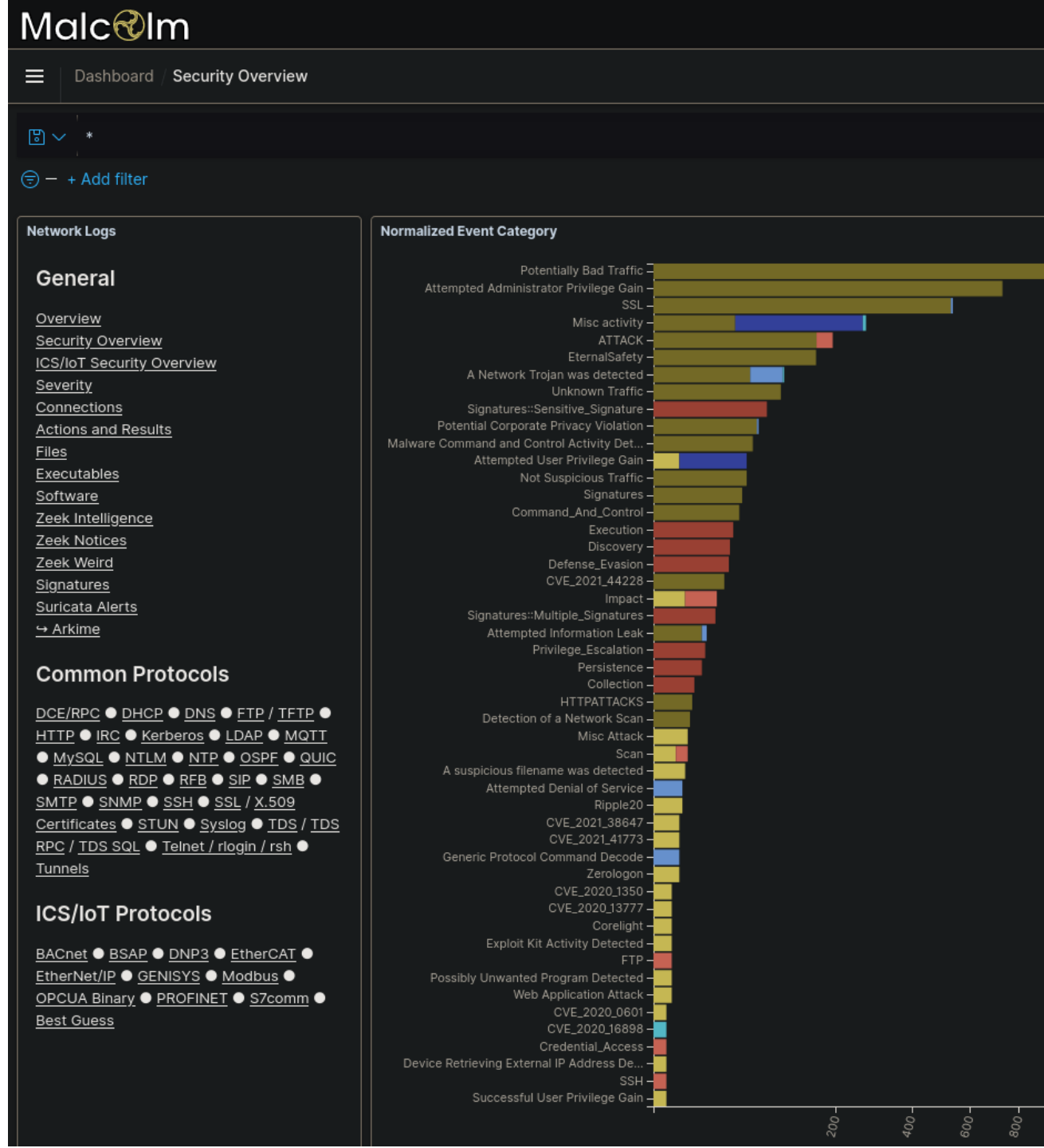
Data Tagging and Enrichment



- Logstash enriches Zeek and Suricata log metadata
 - MAC addresses to hardware vendor
 - GeoIP and ASN lookups
 - Internal/external traffic based on IP ranges
 - Reverse DNS lookups
 - DNS query and hostname entropy analysis
 - Connection fingerprinting (JA3 for TLS, HASSH for SSH, Community ID for flows)
- `tags` field
 - Populated for Arkime sessions, Zeek logs and Arkime alerts with tags provided on upload and words extracted from PCAP filenames
 - `internal_source`, `internal_destination`, `external_source`, `external_destination`, `cross_segment`

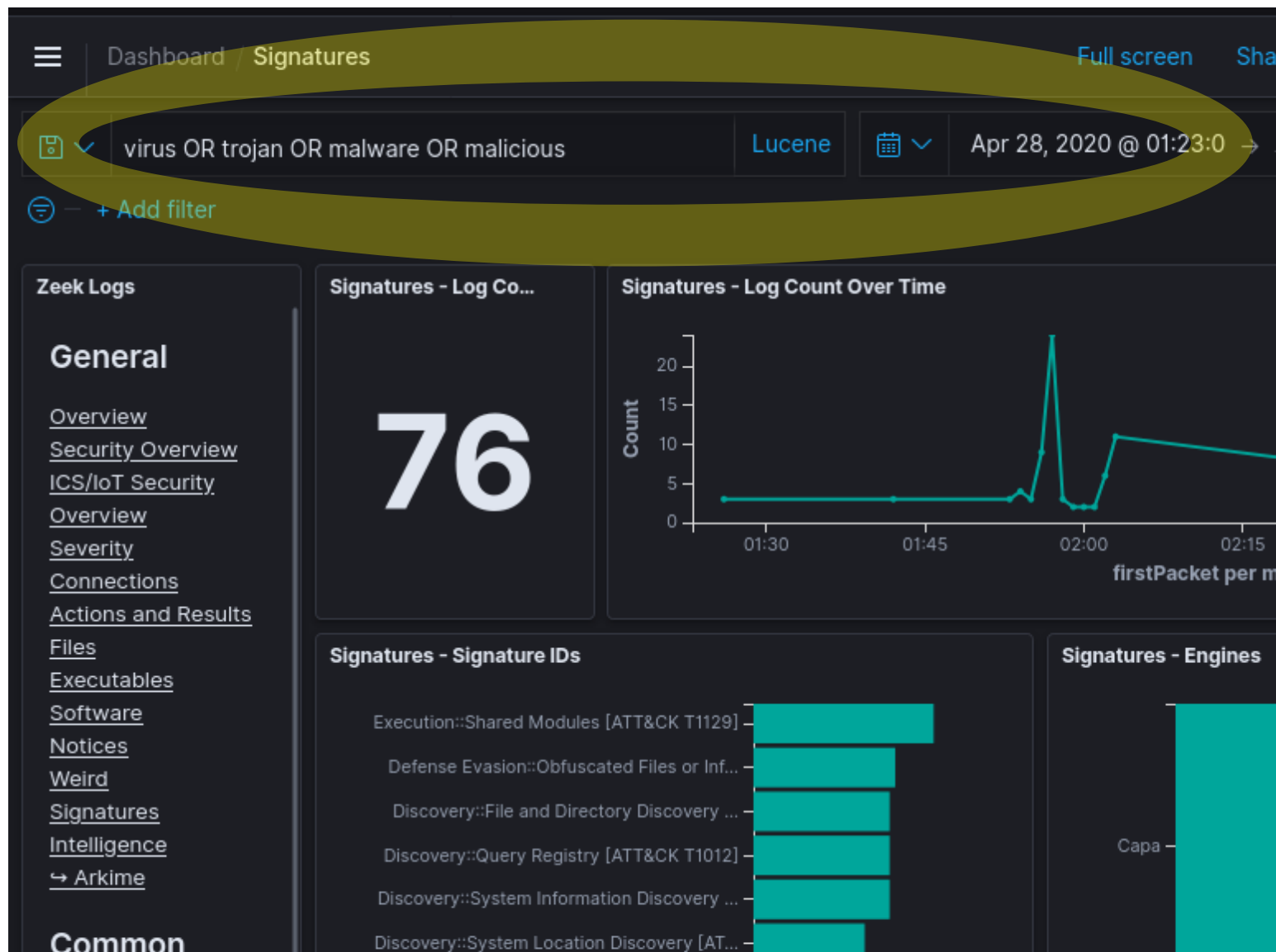


- Front end for Zeek logs and Suricata alerts
- Prebuilt visualizations for all protocols Malcolm parses
- WYSIWYG editors to create custom visualizations and dashboards
- Drill down from high-level trends to specific items of interest
- <https://localhost/dashboards>



Dashboards Filters and Search

- Time filter: define search time frame
- Query bar: write queries in Lucene syntax or DQL (Dashboards Query Language)
- Filter bar: define filters using a UI
 - Pin filters as you move across dashboards
- Save queries and filters for reuse



Overview Dashboards

- High-level view of trends, sessions and events
- Populated from logs across all protocols
- Good jumping-off place for investigation

Network Logs

General

[Overview](#)

[Security Overview](#)

[ICS/IoT Security Overview](#)

[Severity](#)

[Connections](#)

[Actions and Results](#)

[Files](#)

[Executables](#)

[Software](#)

[Zeek Intelligence](#)

[Zeek Notices](#)

[Zeek Weird](#)

[Signatures](#)

[Suricata Alerts](#)

[↔ Arkime](#)

Common Protocols

[DCE/RPC](#) ● [DHCP](#) ● [DNS](#) ● [FTP](#) / [TFTP](#) ●

[HTTP](#) ● [IRC](#) ● [Kerberos](#) ● [LDAP](#) ● [MQTT](#)

● [MySQL](#) ● [NTLM](#) ● [NTP](#) ● [OSPF](#) ● [QUIC](#)

● [RADIUS](#) ● [RDP](#) ● [RFB](#) ● [SIP](#) ● [SMB](#) ●

[SMTP](#) ● [SNMP](#) ● [SSH](#) ● [SSL](#) / [X.509](#)

[Certificates](#) ● [STUN](#) ● [Syslog](#) ● [TDS](#) / [TDS](#)

Normalized Event Categories

Po
Attempted Administr

A Network T

Signatures::
Potential Corpora
Malware Command and Co
Attempted
No

Co

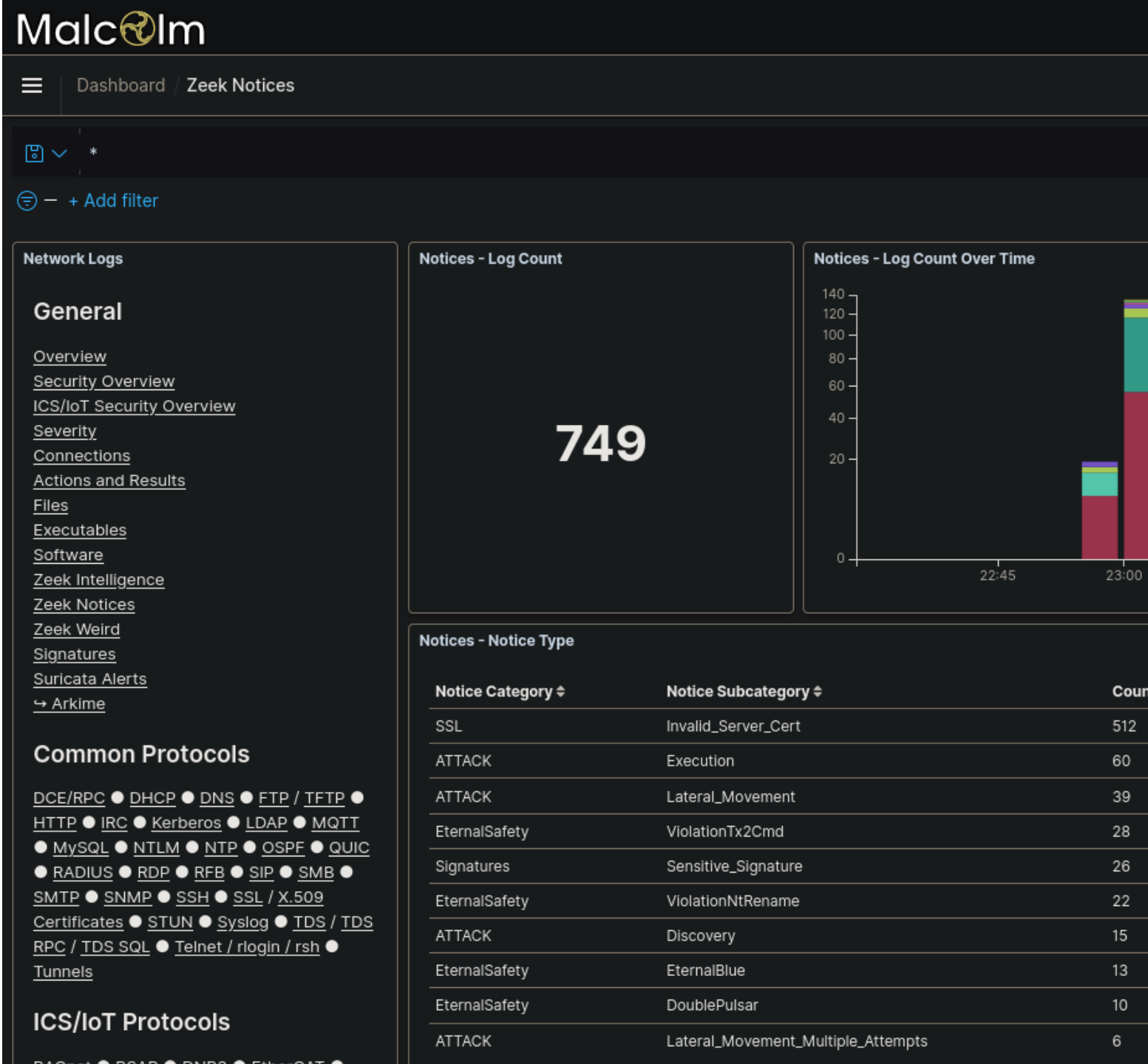
Signatures::
Attempted

Detection

A suspicious file
Attempted

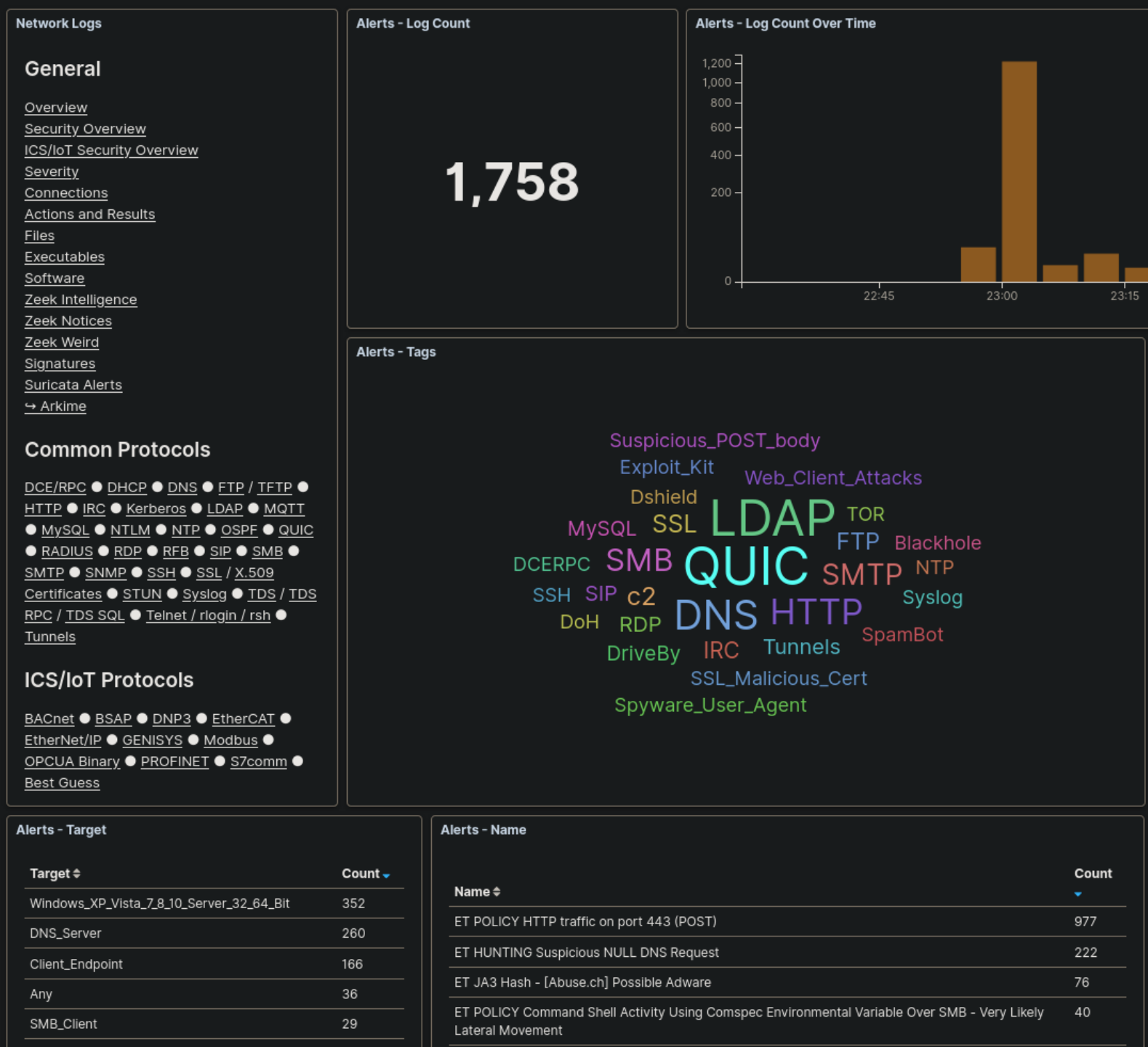
Zeek Notices

- Zeek notices are things that are odd or potentially bad
- In addition to Zeek's defaults, Malcolm raises notices for recent critical vulnerabilities and attack techniques

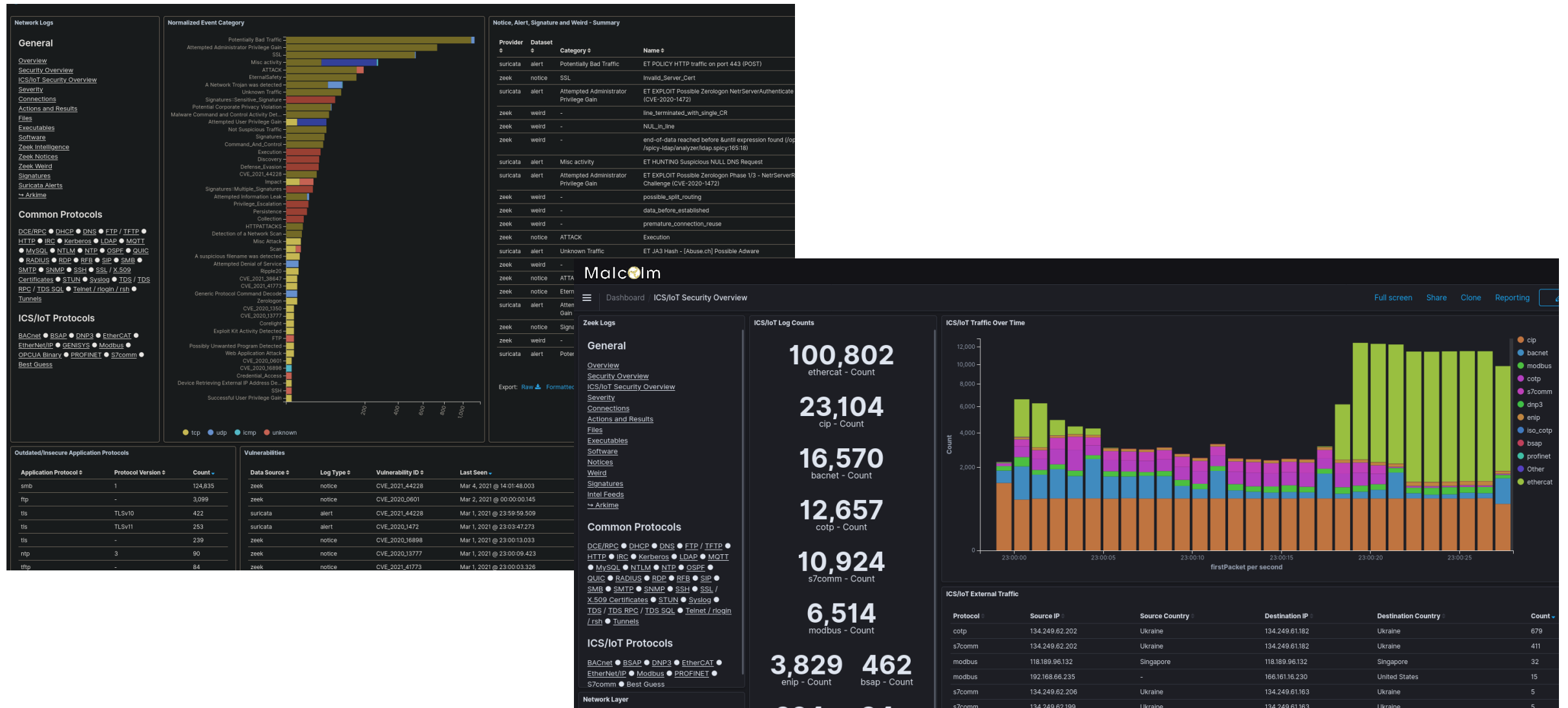


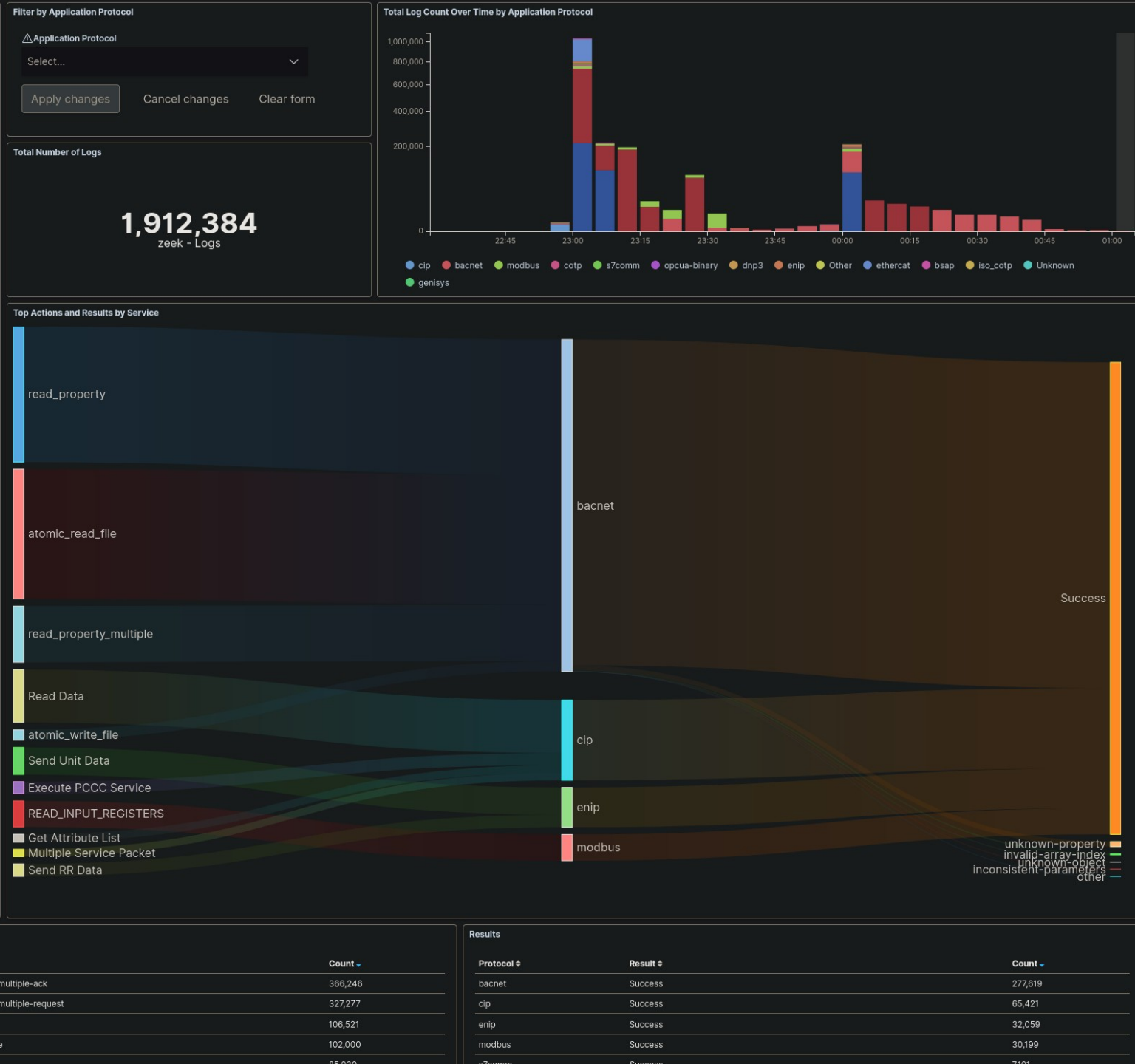
Suricata Alerts

- Protocol-aware Suricata signatures generate alerts for suspect traffic
- Use the default Emerging Threats Open ruleset or custom signatures from other sources



Security & ICS/IoT Security Overviews





Actions and Results

- Malcolm normalizes “action” (e.g., write, read, create file, logon, logoff, etc.) and “result” (e.g., success, failure, access denied, not found) across protocols

Protocol Dashboards

- Highlight application-specific fields of interest
- Grouped by common IT protocols and ICS/IoT protocols
- ICS protocols
 - BACnet
 - BSAP
 - DNP3
 - EtherCAT
 - EtherNet/IP
 - GENISYS
 - Modbus
 - OPCUA Binary
 - PROFINET
 - S7comm

[Zeek Intelligence](#)

[Zeek Notices](#)

[Zeek Weird](#)

[Signatures](#)

[Suricata Alerts](#)

[↔ Arkime](#)

Common Protocols

[DCE/RPC](#) ● [DHCP](#) ● [DNS](#) ● [FTP / TFTP](#) ●
[HTTP](#) ● [IRC](#) ● [Kerberos](#) ● [LDAP](#) ● [MQTT](#)
● [MySQL](#) ● [NTLM](#) ● [NTP](#) ● [OSPF](#) ● [QUIC](#)
● [RADIUS](#) ● [RDP](#) ● [RFB](#) ● [SIP](#) ● [SMB](#) ●
[SMTP](#) ● [SNMP](#) ● [SSH](#) ● [SSL / X.509](#)
[Certificates](#) ● [STUN](#) ● [Syslog](#) ● [TDS / TDS](#)
[RPC / TDS SQL](#) ● [Telnet / rlogin / rsh](#) ●
[Tunnels](#)

ICS/IoT Protocols

[BACnet](#) ● [BSAP](#) ● [DNP3](#) ● [EtherCAT](#) ●
[EtherNet/IP](#) ● [GENISYS](#) ● [Modbus](#) ●
[OPCUA Binary](#) ● [PROFINET](#) ● [S7comm](#) ●
[Best Guess](#)

Notices - Notice Type

Notice Category

SSL

ATTACK

ATTACK

EternalSafety

Signatures

EternalSafety

ATTACK

EternalSafety

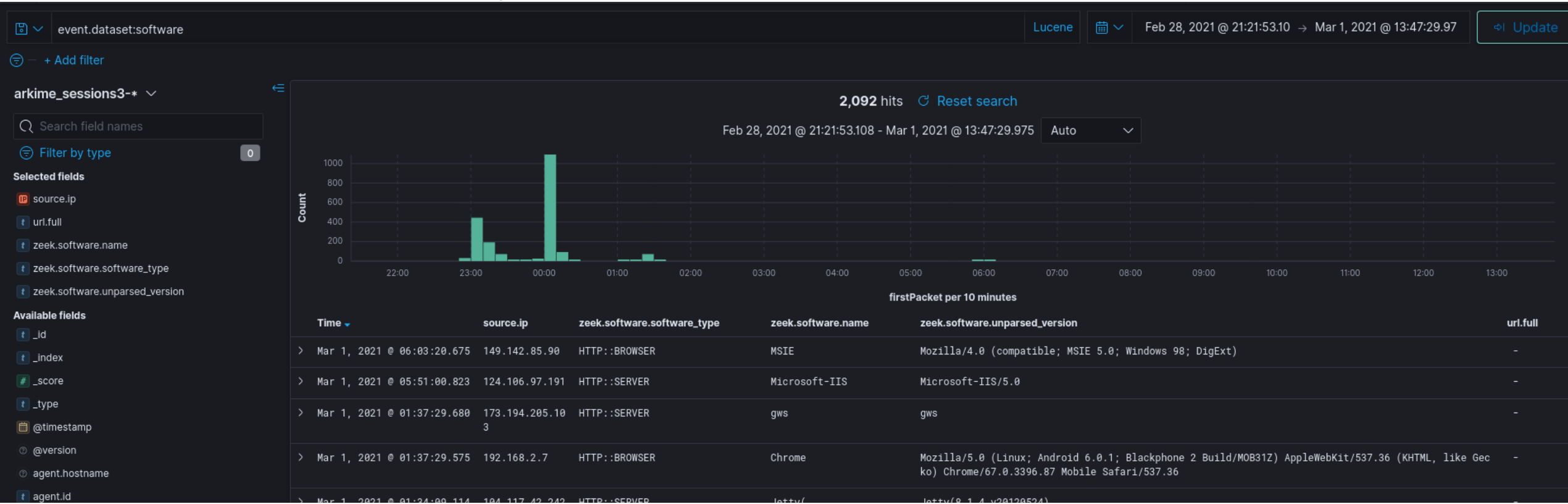
EternalSafety

ATTACK

Export: [Raw](#) [Download](#) [Format](#)

Discover

- Field-level details of logs matching filter criteria
- Create and view saved searches and column configurations
- View other events just before and after an event



New Visualization

Q Filter



Area



Controls



Coordinate
Map



Data Table



Gantt Chart



Gauge



Goal



Heat Map



Horizontal Bar



Line



Markdown



Metric



Pie



Region Map



Sankey
Diagram



TSVB



Tag Cloud



Timeline



Vega



Vertical Bar

Custom Visualizations

- Create new visualizations from scratch or based on existing charts or dashboards

Search Syntax Comparison

	Arkime	Dashboards (Lucene)	Dashboards (DQL)
Field exists	<code>event.dataset == EXISTS!</code>	<code>_exists_:event.dataset</code>	<code>event.dataset:*</code>
Field does not exist	<code>event.dataset != EXISTS!</code>	<code>NOT _exists_:event.dataset</code>	<code>NOT event.dataset:*</code>
Field matches a value	<code>port.dst == 22</code>	<code>destination.port:22</code>	<code>destination.port:22</code>
Field does not match a value	<code>port.dst != 22</code>	<code>NOT destination.port:22</code>	<code>NOT destination.port:22</code>
Field matches at least one of a list of values	<code>tags == [external_source, external_destination]</code>	<code>tags:(external_source OR external_destination)</code>	<code>tags:(external_source or external_destination)</code>
Field range (inclusive)	<code>http.statuscode >= 200 && http.statuscode <= 300</code>	<code>http.statuscode:[200 TO 300]</code>	<code>http.statuscode >= 200 and http.statuscode <= 300</code>

Search Syntax Comparison (cont.)

	Arkime	Dashboards (Lucene)	Dashboards (DQL)
Field range (exclusive)	<code>http.statuscode > 200 && http.statuscode < 300</code>	<code>http.statuscode:{200 TO 300}</code>	<code>http.statuscode > 200 and http.statuscode < 300</code>
Field range (mixed exclusivity)	<code>http.statuscode >= 200 && http.statuscode < 300</code>	<code>http.statuscode:[200 TO 300}</code>	<code>http.statuscode >= 200 and http.statuscode < 300</code>
Match all search terms (AND)	<code>(tags == [external_source, external_destination]) && (http.statuscode == 401)</code>	<code>tags:(external_source OR external_destination) AND http.statuscode:401</code>	<code>tags:(external_source or external_destination) and http.statuscode:401</code>
Match any search terms (OR)	<code>(zeek_ftp.password == EXISTS!) (zeek_http.password == EXISTS!) (zeek.user == "anonymous")</code>	<code>_exists_:zeek_ftp.password OR _exists_:zeek_http.password OR zeek.user:"anonymous"</code>	<code>zeek_ftp.password:* or zeek_http.password:* or zeek.user:"anonymous"</code>

Search Syntax Comparison (cont.)

	Arkime	Dashboards (Lucene)	Dashboards (DQL)
Global string search (anywhere in the document)	all Arkime search expressions are field-based	microsoft	microsoft
Wildcards	host.dns == "*micro?oft*" (? for single character, * for any characters)	dns.host:*micro?oft* (? for single character, * for any characters)	dns.host:*micro*ft* (* for any characters)
Regex	host.http == /. *www\.f.*k\.com.* /	zeek_http.host: /. *www\.f.*k\.com.* /	Dashboards Query Language does not currently support regex
IPv4 values	ip == 0.0.0.0/0	source.ip:"0.0.0.0/0" OR destination.ip:"0.0.0.0/0"	source.ip:"0.0.0.0/0" OR destination.ip:"0.0.0.0/0"
IPv6 values	(ip.src == EXISTS! ip.dst == EXISTS!) && (ip != 0.0.0.0/0)	(_exists_:source.ip AND NOT source.ip:"0.0.0.0/0") OR (_exists_:destination.ip AND NOT destination.ip:"0.0.0.0/0")	(source.ip:* and not source.ip:"0.0.0.0/0") or (destination.ip:* and not destination.ip:"0.0.0.0/0")

Search Syntax Comparison (cont.)

	Arkime	Dashboards (Lucene)	Dashboards (DQL)
GeoIP information available	<code>country == EXISTS!</code>	<code>_exists_:destination.geo OR _exists_:source.geo</code>	<code>destination.geo:* or source.geo:*</code>
Log type	<code>event.dataset == notice</code>	<code>event.dataset:notice</code>	<code>event.dataset:notice</code>
IP CIDR Subnets	<code>ip.src == 172.16.0.0/12</code>	<code>source.ip:"172.16.0.0/12"</code>	<code>source.ip:"172.16.0.0/12"</code>
Search time frame	Use Arkime time bounding controls under the search bar	Use Dashboards time range controls in the upper right-hand corner	Use Dashboards time range controls in the upper right-hand corner
GeoIP information available	<code>country == EXISTS!</code>	<code>_exists_:destination.geo OR _exists_:source.geo</code>	<code>destination.geo:* or source.geo:*</code>

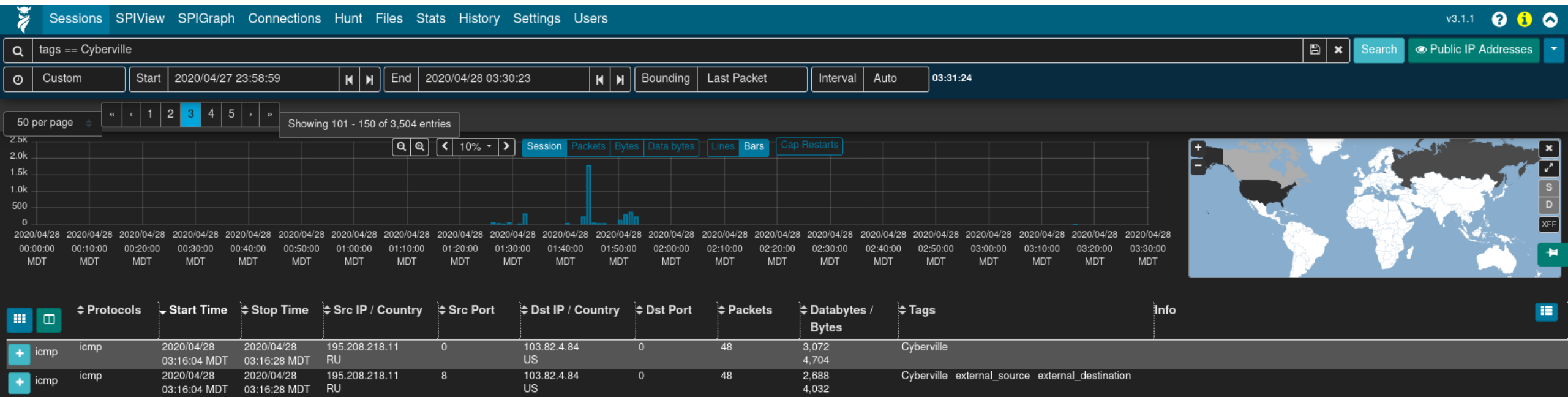
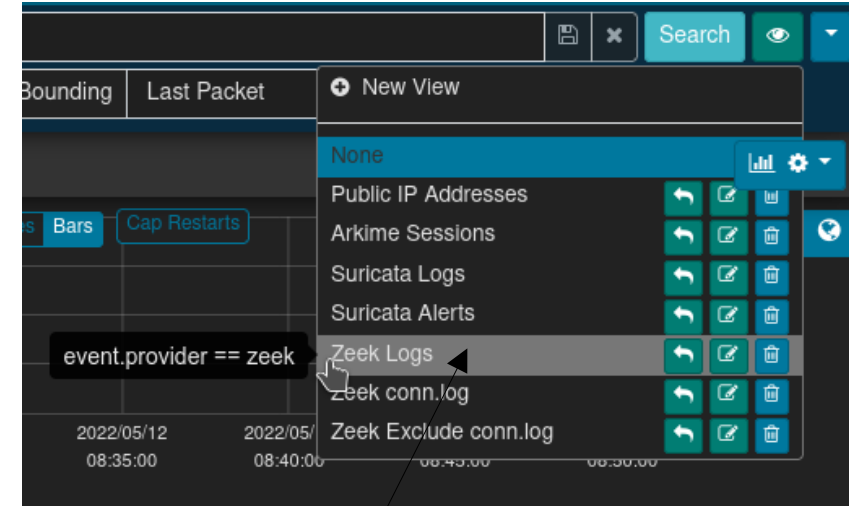


Arkime

- Front end for **both** enriched Zeek logs, Suricata alerts and Arkime sessions
 - Malcolm's custom Arkime Zeek data source adds full support for Zeek logs to Arkime, including ICS protocols
- Filter by data source (Zeek, Suricata or Arkime); or, view together
- “Wireshark at scale”: full PCAP availability for
 - viewing packet payload
 - exporting filtered and joined PCAP sessions
 - running deep-packet searches
- <https://localhost>

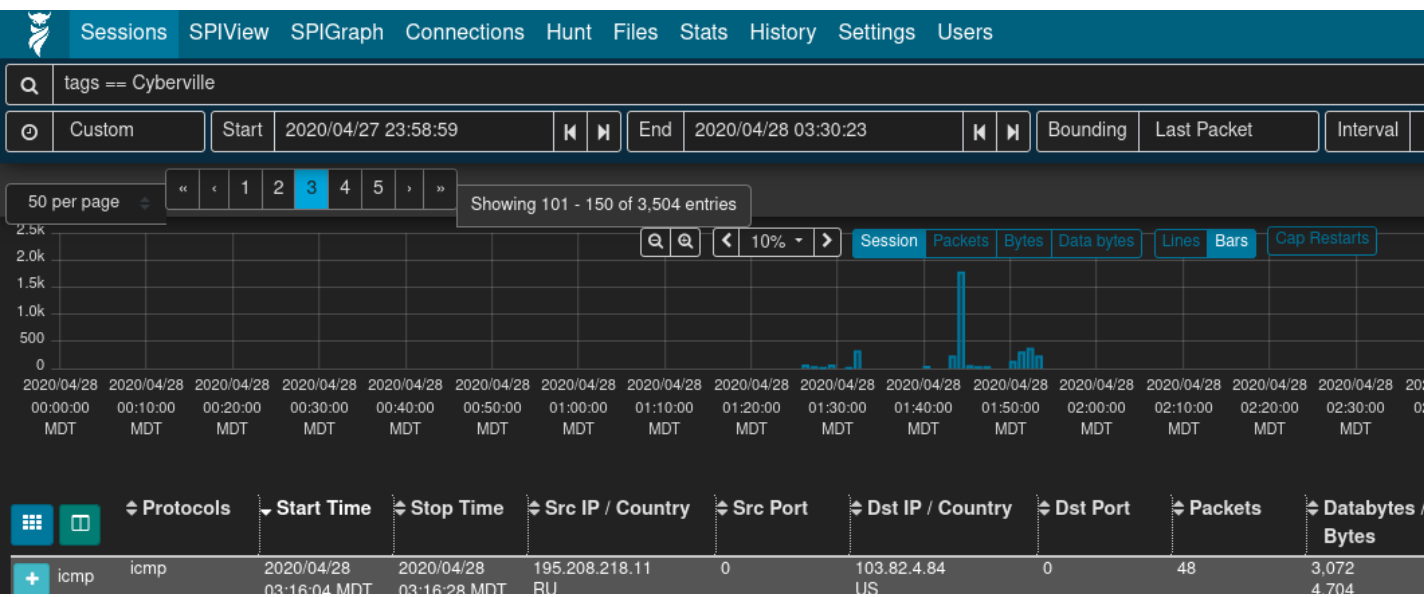
Arkime Filters and Search

- Time filter: define search time frame
- Map filter: restrict results to geolocation
- Query bar: write queries in Arkime syntax
- Views: overlay previously-specified filters on current search



Sessions

- Field-level details of sessions/logs matching filters
- Similar to Dashboards' Discover



The screenshot shows the 'Zeek http.log' interface with a search filter 'protocols == http && tags == external_destination'. The time range is from 2020/11/11 06:23:48 to 2021/05/30 06:00:53. The interface displays detailed session information for a specific log entry.

Log Type: http
Malcolm Data Source: zeek
Malcolm Node: filebeat
Originating Host: 217.226.31.170
Originating GeoIP Country: Germany
Originating GeoIP City: Bremen
Responding Host: 124.106.97.191
Responding GeoIP Country: Philippines
Responding GeoIP City: Santa Elena
Originating Port: 4230
Responding Port: 80
Related IP: 217.226.31.170 124.106.97.191
Protocol: tcp
Service: http
Service Version: 1.1
Action: GET
Result: Bad Gateway
Severity: 20
Risk Score: 20
Severity Tags: External traffic
File Magic: text/html

Pipeline Depth: 1
Request Method: GET
URI: /_vti_bin/.../winnt/system32/cmd.exe?/c+dir+x:\\c+dir+x:\\c+dir+x:\\
Version: 1.1

Packet Payloads

- Displayed for Arkime sessions with full PCAP (i.e., not Zeek logs)
- File carving on the fly
- Download session PCAP
- Examine payload with CyberChef

Source	Destination
<pre>GET /PostExploitation/PCAnyPass.exe HTTP/1.1 Accept: text/html, application/xhtml+xml, */* Referer: http://10.10.10.11/PostExploitation/ Accept-Language: en-US User-Agent: Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; Trident/5.0) Accept-Encoding: gzip, deflate Host: 10.10.10.11 Connection: Keep-Alive</pre>	<pre>HTTP/1.0 200 OK Server: SimpleHTTP/0.6 Python/2.7.17 Date: Fri, 17 Apr 2020 19:21:32 GMT Content-type: application/x-msdos-program Content-Length: 49152 Last-Modified: Fri, 16 Apr 2010 19:09:50 GMT PCAnyPass.exe</pre>

Packets200

naturalasciitf8hex

Show Packets

Line Numbers

Uncompress

Show Image & Files

Show Info

File Bytes:

base64

CyberChef

Export PCAP

- Creates a new PCAP file from filtered sessions
- Include open, visible or all matching sessions
- Apply “Arkime Sessions” view to sessions first
- Narrow as much as possible prior to exporting (huge PCAP files are a pain)

The screenshot displays the Arkime Sessions web interface. At the top, a navigation bar includes tabs for Sessions, SPIView, SPIGraph, Connections, Hunt, Files, Stats, History, Settings, and Users. A search bar contains the query 'country != US && protocols == http'. Below this, a filter bar shows 'Custom' filter, start/end times (2021/02/28 23:59:11 to 2021/03/01 00:28:26), and various options like Bounding, Last Packet, Interval, and Auto. The 'Export PCAP' button is visible on the right. The main area shows a list of sessions with columns for Session, Packets, Bytes, Data bytes, Lines, Bars, and Cap Restarts. A bar chart at the bottom left shows session activity over time. On the right, a world map highlights the source countries of the sessions. The bottom status bar shows details for a selected session: tcp http, 2021/03/01, 10.0.52.164, 2550, 61.8.0.17, 80, 7,195, 5,160,414, HTTP, out-of-order-dst, and the URI 'mirror.pacific.net.au/openoffice/stable/2.0.0/OOo_2.0.0_Win32Intel_install.exe'.

Sessions SPIView SPIGraph Connections Hunt Files Stats History Settings Users v3.1.1

country != US && protocols == http Search Arkime Sessions

Custom Start 2021/02/28 23:59:11 End 2021/03/01 00:28:26 Bounding Last Packet Interval Auto 00:29:15

Open Items Visible Items Matching Items Include same time period linked segments (slow) Filename US_HTTP.pcap Export PCAP

50 per page 1 2 3 Showing 1 - 50 of 120 entries

15 10 5 0 2021/03/01 00:00:00 2021/03/01 00:02:00 2021/03/01 00:04:00 2021/03/01 00:06:00 2021/03/01 00:08:00 2021/03/01 00:10:00 2021/03/01 00:12:00 2021/03/01 00:14:00 2021/03/01 00:16:00 2021/03/01 00:18:00 2021/03/01 00:20:00 2021/03/01 00:22:00 2021/03/01 00:24:00 2021/03/01 00:26:00 2021/03/01 00:28:00

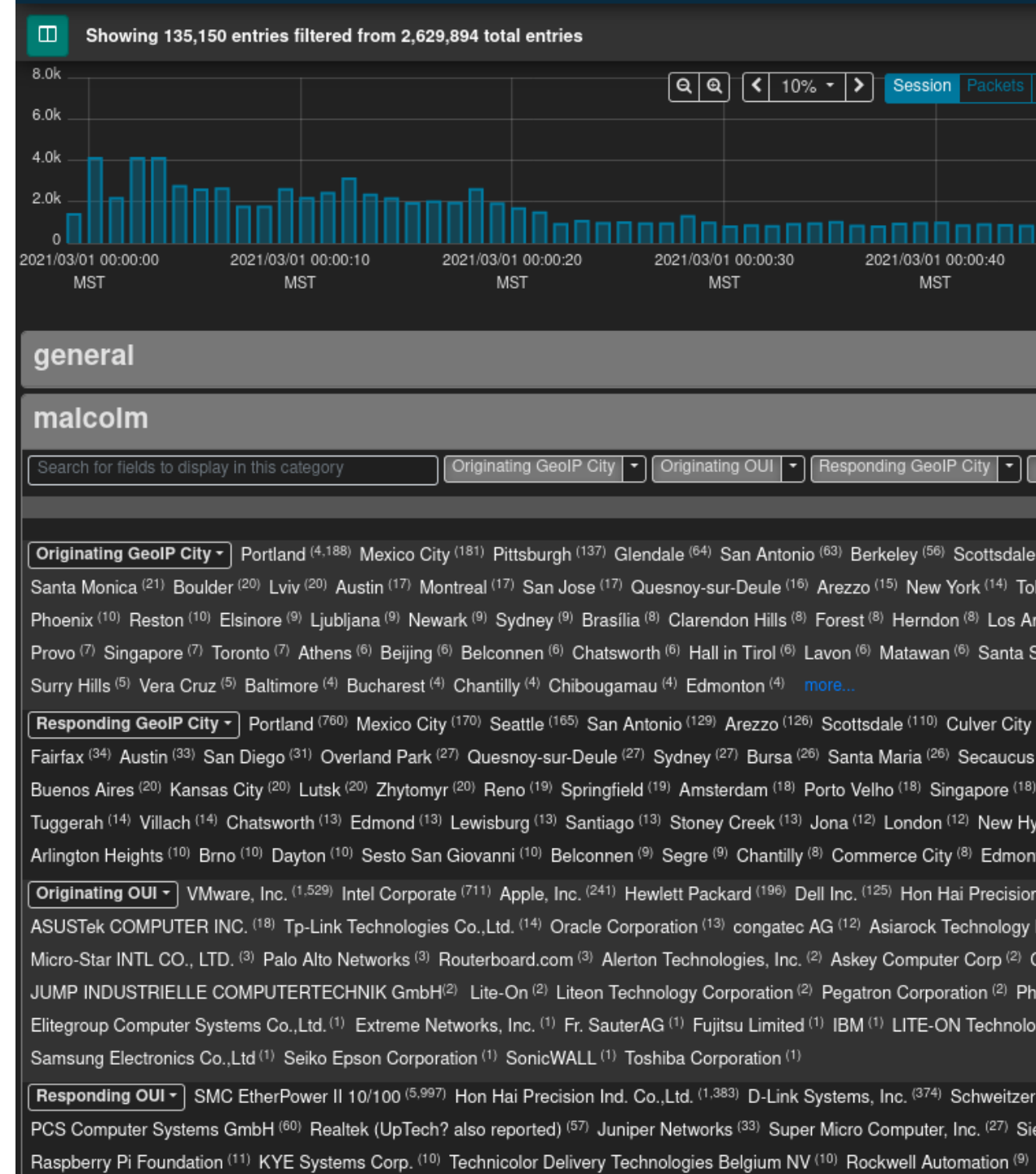
Session Packets Bytes Data bytes Lines Bars Cap Restarts

Protocols Start Time Stop Time Src IP / Country Src Port Dst IP / Country Dst Port Packets Databytes / Bytes Tags Info

tcp http 2021/03/01 2021/03/01 10.0.52.164 2550 61.8.0.17 80 7,195 5,160,414 HTTP out-of-order-dst URI mirror.pacific.net.au/openoffice/stable/2.0.0/OOo_2.0.0_Win32Intel_install.exe

SPIView

- Explore “top n ” and field cardinality for all fields of both Arkime sessions and Zeek logs
- Apply filters or pivot to Sessions or SPIGraph view for field values of interest
- Limit search to ≤ 1 week before using (it runs many queries)



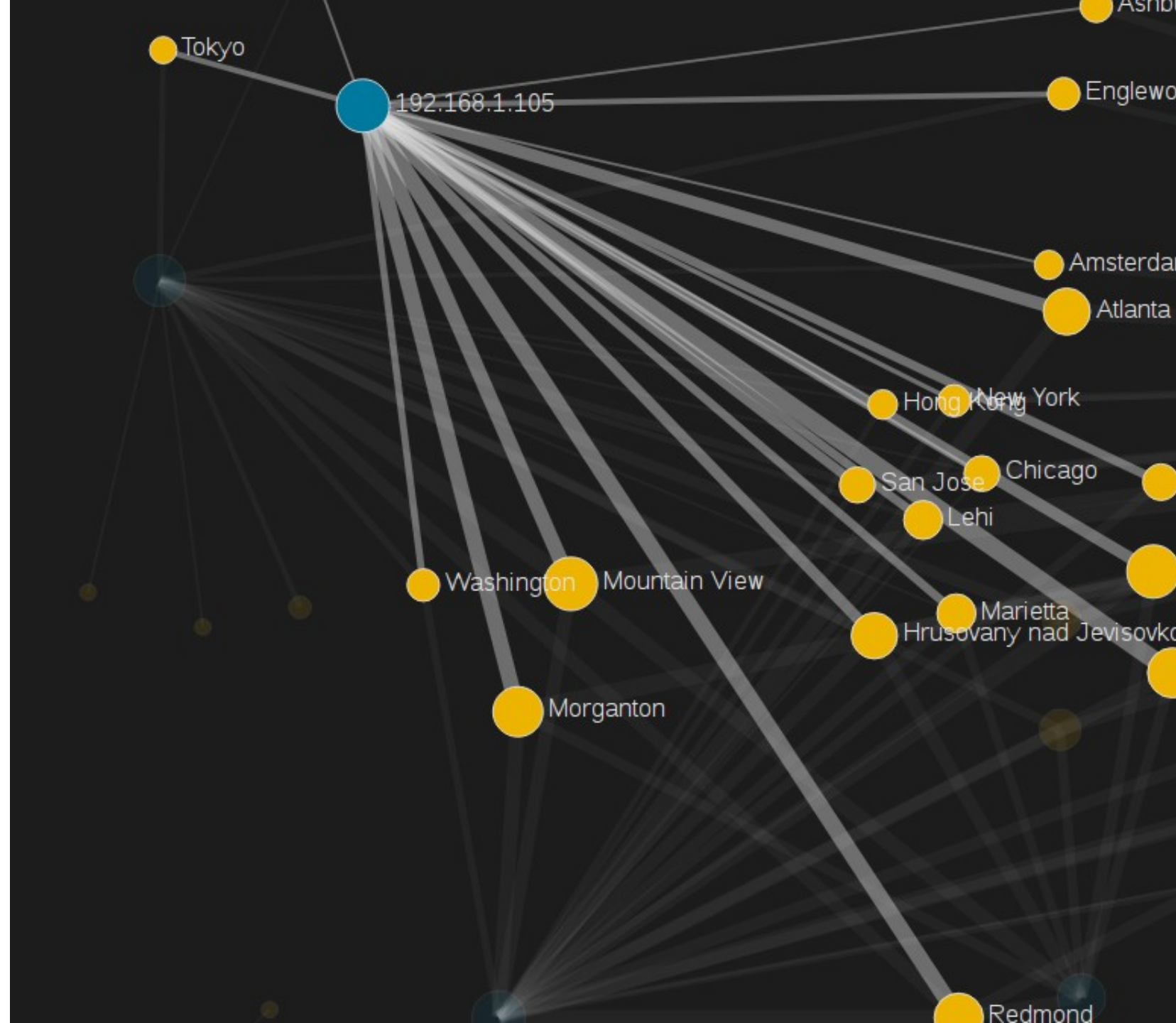
SPIGraph

- View “top n ” field values chronologically and geographically
- Identify trends and patterns in network traffic



Connections

- Visualize logical relationship between hosts
- Use any combination of fields for source and destination nodes
- Compare current vs. previous (baseline) traffic



Packet Search (“Hunt”)

- Deep-packet search (“PCAP grep”) of session payloads
- Search for ASCII, hex codes or regular expression matches
- Apply “Arkime Sessions” view to sessions first

Sessions

SPIView

SPIGraph

Connections

Hunt

Files

Stats

History

Settings

Users

v3.1.1

protocols == http

Search

Arkime Sessions

All (careful)

Start 1969/12/31 17:00:00

End 2021/12/06 12:10:02

Bounding

Last Packet

Creating a new packet search job will search the packets of 2,906 sessions.

Create a packet search job

Hunt Job History

Search your packet search job history

50 per page

<< < 1 > >>

Showing 1 - 1 of 1 entries

Status	Matches	Name	User	Search text	Notify	Created	ID
<div><div>x</div><div>✓</div><div>100%</div></div>	141	HTTP with password		password (ascii)		2021/12/06 12:12:27 MST	s5YpkX0BTA40FhD4X7dA

✓ This hunt is finished

👁 Found 141 sessions matching password (ascii) of 2,908 sessions searched

🕒 Created: 2021/12/06 12:12:27 MST

🕒 Last Updated: 2021/12/06 12:12:32 MST

🔍 Examining 500 raw source and destination packets per session

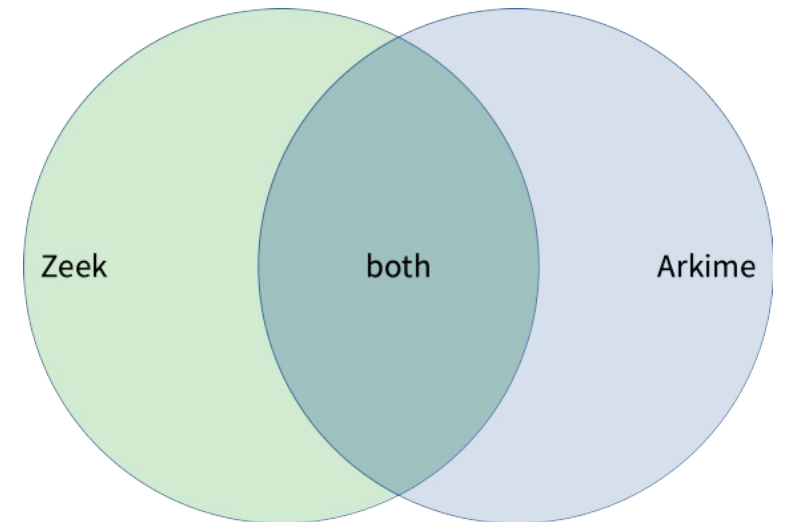
🔍 The sessions query expression was: protocols == http

🔍 The sessions query view was: Arkime Sessions

🕒 The sessions query time range was from 1969/12/31 17:00:00 MST to 2021/12/06 12:10:02 MST

Data Source Correlation

- Search syntax is different between Arkime and Dashboards (and in some cases, so are field names)
 - See search syntax comparison table, Malcolm and Arkime docs
- Despite considerable overlap, there are differences in protocol parser support among Zeek, Suricata and Arkime
 - Learning the strengths of each will help you more effectively find the good stuff



Correlate Zeek or Suricata Logs and Packet Payloads

- Correlate Zeek or Suricata logs and Arkime sessions using common fields
- `communityId` fingerprints flows to bridge data sources
- `rootId/event.id` filters logs for the same session
- Filter community ID OR'ed with `event.id` to see all Arkime sessions and Zeek or Suricata logs for the same traffic

```
communityId == "1:deE1x8qWJLk/Ul6TKoVasnf4FZw=" || event.id == "C2EHCo28ylyUYz82Wk || event.id == 1553655336047566"
```

communityId == "1:deE1x8qWJLk/Ul6TKoVasnf4FZw=" event.id == C2EHCo28ylyUYz82Wk event.id == 1553655336047566											
Custom	Start	2021/03/01 00:06:01	End	2021/03/01 00:06:58	Bounding	Last Packet	Interval	Auto	00:00:57		
50 per page 1 Showing 1 - 6 of 6 entries											
	Protocols	Data Source	Log Type	Start Time	Stop Time	Src IP / Country	Dst IP / Country	Dst Port	Packets	Databytes / Bytes	Info
+	http http	zeek	files	2021/03/01 00:06:16	2021/03/01 00:06:16	204.152.184.134 US	10.0.52.164			6,550,964 78,597,807	Severity Tags File transfer (high concern)
+	tcp tcp http	zeek	http	2021/03/01 00:06:16	2021/03/01 00:06:16	10.0.52.164	204.152.184.134 US	80		6,550,964	URI mirrors.isc.org/pub/openoffice/stable/2.0.0/OOo_2.0.0_Win32Intel_install.exe Severity Tags Outbound traffic File transfer (high concern)
+	tcp tcp http	arkime	session	2021/03/01 00:06:16	2021/03/01 00:06:39	10.0.52.164	204.152.184.134 US	80	9,354	6,551,694 7,057,419	URI mirrors.isc.org/pub/openoffice/stable/2.0.0/OOo_2.0.0_Win32Intel_install.exe
+	tcp tcp http	suricata	alert	2021/03/01 00:06:16	2021/03/01 00:06:17	204.152.184.134 US	10.0.52.164	2646	58	47,435 47,435	Severity Tags Suricata Alert Inbound traffic Service on non-standard port
+	tcp tcp http	zeek	conn	2021/03/01 00:06:16	2021/03/01 00:06:39	10.0.52.164	204.152.184.134 US	80	9,354	6,551,694 6,926,454	Severity Tags Outbound traffic
+		zeek	signatures	2021/03/01 00:06:16	2021/03/01 00:06:16						Severity Tags Signature

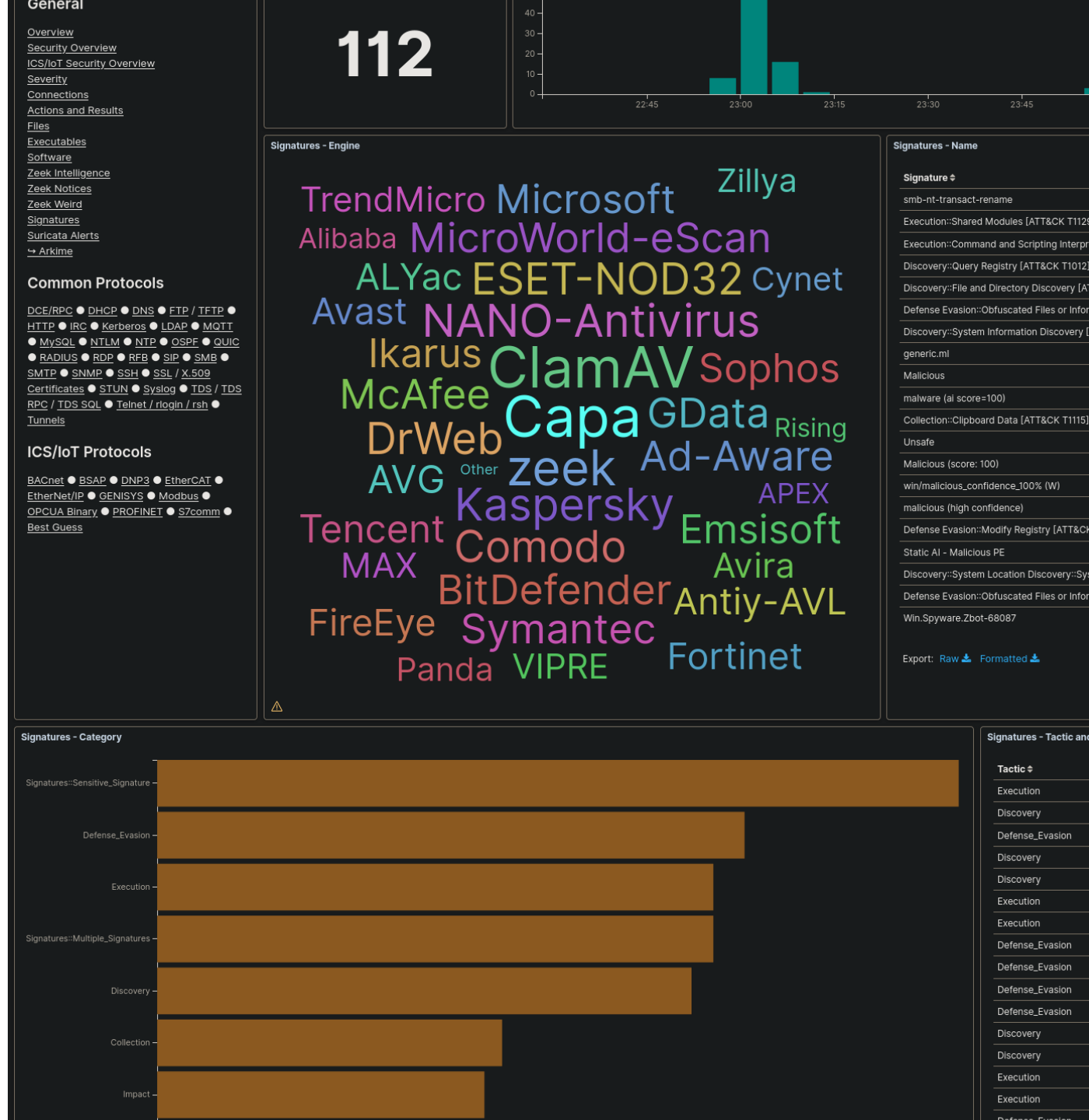
File Analysis

- Zeek can “carve” file transfers from common protocols
- Malcolm can examine carved files and flag hits
 - ClamAV – open source antivirus engine
 - YARA – pattern matching swiss army knife
 - Capa – portable executable capabilities analyzer
 - VirusTotal – online database of file hashes
 - requires API token and internet connection
- Triggering files can be saved to `zeek-logs/extract_files` under Malcolm directory for further analysis
 - Be careful! Carved files may contain live malware!



Signatures

- Signatures dashboard in Dashboards shows scanned file hits
- Use `zeek.fuid` field in *Signatures - Logs* table to pivot to connection UID (`zeek.uid`) and other logs with pertinent session details



Search Tips

- Always check your search time frame
- “Zoom in” (apply filters) for a particular field value, pivot to another field then “zoom out” (remove filters)
- Most UI controls can work with any data field (2000+)
- Filter on `event.dataset` (e.g., `conn` to see `conn.log`)
- Filter on protocol regardless of data source (e.g., `protocol:http` in Dashboards and `protocols == http` in Arkime)
- Use tags

Malcolm



Thank you!

Visit [Malcolm on GitHub](#) to read the docs, make suggestions, report issues and st★r to show your support!

Malcolm is Copyright © 2022 Battelle Energy Alliance, LLC, and is developed and released as open-source software through the cooperation of the Cybersecurity and Infrastructure Security Agency of the US Department of Homeland Security.