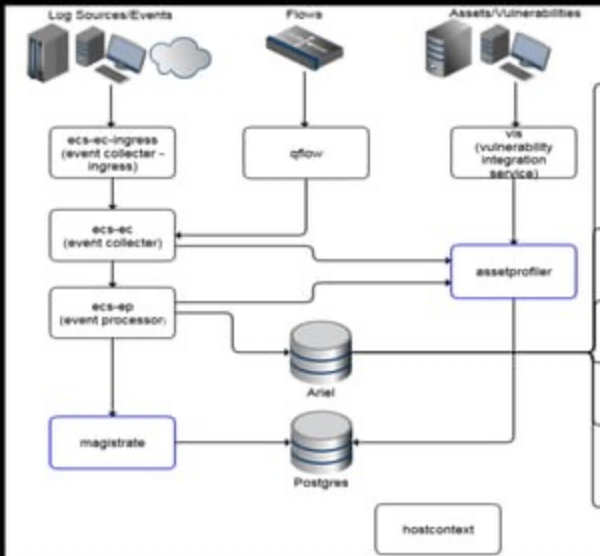


# How QRadar SIEM Collected Security Data

# QRadar Data Flow - Overall



# From an Appliance Perspective

## Event Collector Capabilities



# From an Appliance Perspective

## Event/Flow Processor Capabilities

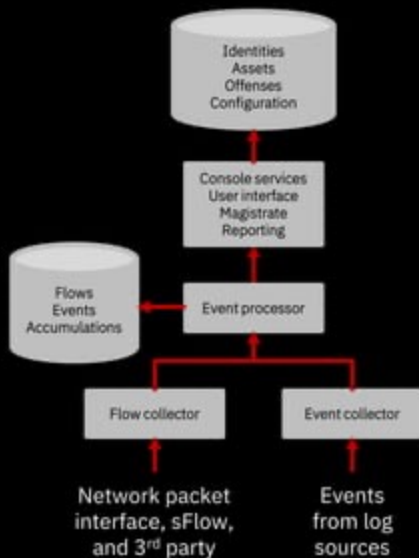


# From an Appliance Perspective

**AIO/Console Capabilities**



# High-level component architecture a



Flow and  
**database**

- If accum
- in Ariel a

- As soon
- proof)

- Data can

Offenses,  
stored in  
Console

- Provides
- process

Secure SS  
a distribu

# QRadar Data Flow - Overall

# Collecting and Normalizing raw events

**An event** is a record from a device that describes an incident.

QRadar SIEM normalizes the varied information found in events.

- Normalizing means to map information to common fields.
  - SRC\_IP, Source, IP, and others are normalized to SRC\_IP.
  - user\_name, username, login, and others are normalized to user\_name.
- Normalized events are mapped to high-level and low-level processing.
- After raw events are normalized, it is easy to search for normalized events.



# Event data pipeline

Event data is sent to or pulled by QRadar

**Event Collector Ingress** – Responsible for collecting data at all times (zero event loss)

Data is collected and buffered during patch and deploys and processed once the operation is complete

**Protocols** – Reads or pulls raw data from network devices (e.g: Windows Servers, Firewalls, etc)

**Throttle Filter - Licensing** - On a second-by-second basis, slows down the incoming rate so it does not exceed the license on the appliance.

Events are sent to ecs-ec-parse to be parsed

# Event data pipeline

Event data is received from the ecs-ec-ingress

**Parsing** – DSMs / LSX / CEP – take the raw data and normalize it into a common structure.

**Coalescing** - “Event Compression”. Find nearly identical events and delete one and increase the event count on the record. Key is: source IP, dest IP, dest port, QID, username

**Forwarding** - Applies routing rules for the system, such as sending event data to offsite targets, external Syslog systems, JSON systems, and other SIEMs.

**Log Only/Data Store** supports the storage of an unlimited number of logs without counting against the EPS License

Events are then sent to the **Event Processor** component and pass through the Custom Rules Engine (CRE).

# Events not counted against the EPS l

- The list of log source types that do not incur EPS hits are
  - System Notification
  - Custom Rule Engine (CRE)
  - SIM Audit
  - Anomaly Detection Engine
  - Asset Profiler
  - Search Results from scheduled searches
  - Health Metrics
  - Sense DSM
  - Risk Manager questions, Simulations and internal l
- Log Only/Data Store
- Supports the storage of an unlimited number of log  
SIEM license
- Enables an organization to build custom apps and m  
deeper insights into IT environments.

# Event Coalescing

- Event Coalescing is a method of reducing the data
- As data arrives in the pipeline QRadar will attempt single event.
- Coalescing occurs after licensing and parsing
- Coalescing is indexed by Log Source, QID, Source and Username.
- If more than 4 events arrive within a 10 second window identical any additional events beyond the 4<sup>th</sup> will
- Coalesced events can be identified by looking at the viewer, if the Event Count is >1 the event has been
- Coalescing can be turned on or off per log source on the system setting page.

# QRadar Data Flow - Overall



# Flow collection and processing

A *flow* is a communication session between two hosts

QFlow Collectors read packets from the wire or receive them from other collectors

QFlow Collectors convert all gathered network data into flows; they include such details as:

- when, who, how much, protocols, and options.

Flow Type	First Packet Time	Source IP	Source Port	Destination IP	Destination Port	Protocol	Misc. details
<input type="checkbox"/>	Oct 14, 2014, 7:00:13 AM	192.168.0.1	61190	202.12.27.33	53	udp_ip	Misc. domain
<input type="checkbox"/>	Oct 14, 2014, 6:58:59 AM	192.168.0.1	64334	192.168.10.10	22	tcp_ip	RemoteAccess
<input type="checkbox"/>	Oct 14, 2014, 7:00:53 AM	0.0.0.0	548	0.0.0.0	547	udp_ip	Other
<input type="checkbox"/>	Oct 14, 2014, 6:58:59 AM	192.168.0.1	64334	192.168.10.10	22	tcp_ip	RemoteAccess
<input type="checkbox"/>	Oct 14, 2014, 6:58:59 AM	192.168.0.1	64334	192.168.10.10	22	tcp_ip	RemoteAccess
<input type="checkbox"/>	Oct 14, 2014, 7:00:09 AM	192.168.0.1	61190	192.203.230.18	53	udp_ip	Misc. domain
<input type="checkbox"/>	Oct 14, 2014, 7:00:53 AM	0.0.0.0	548	0.0.0.0	547	udp_ip	Other
<input type="checkbox"/>	Oct 14, 2014, 7:00:24 AM	192.168.0.1	64348	192.168.10.10	443	tcp_ip	WebSecure
<input type="checkbox"/>	Oct 14, 2014, 7:00:05 AM	192.168.0.1	61708	192.168.10.1	53	udp_ip	Misc. domain
<input type="checkbox"/>	Oct 14, 2014, 6:58:59 AM	192.168.0.1	61897	192.168.99.1	53	udp_ip	Misc. domain
<input type="checkbox"/>	Oct 14, 2014, 7:00:01 AM	192.168.0.1	64335	192.168.10.10	443	tcp_ip	WebSecure
<input type="checkbox"/>	Oct 14, 2014, 7:00:05 AM	192.168.0.1	N/A	192.168.10.12	548	icmp_ip	ICMP Dest

# Flow pipeline

The **QFlow** component collects and creates flow information from internal and external flow sources

**Event Collector** – Responsible for parsing and normalizing incoming flows

**Asymmetric recombination** - Responsible for combining two sides of each flow when data is provided asymmetrically

**Deduplication** - Flow deduplication is a process that removes duplicate flows when multiple Flow Collectors provide data to Flow Processors appliances.

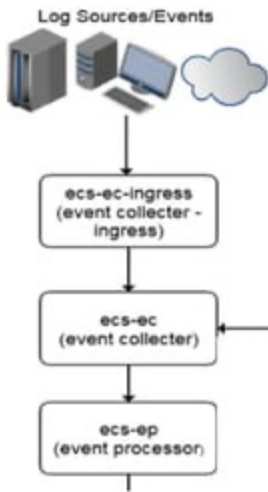
**Flow Governor** - Monitors the number of incoming flows to the system to manage input queues and licensing.

**Custom flow properties** – extracts any properties defined in the Custom Flow Properties

**Forwarding** - Applies routing rules for the system, such as sending flow data to offsite targets, external Syslog systems, JSON systems, and other SIEMs.

Flows are then sent to the **Event Processor** component and pass through the Custom Rules Engine (CRE). They are tested and correlated against the rules that are configured

# QRadar Data Flow - Overall





# Event & Flow Correlation and Proces

After Events and Flows are normalized they are then sent to the Event Processor for processing

Licensing is applied again on ingress to the EP

The CRE or Custom Rules Engine Applies the correlation rules that were created in the UI.

Flow data is then sent to the Ariel Database for storage.

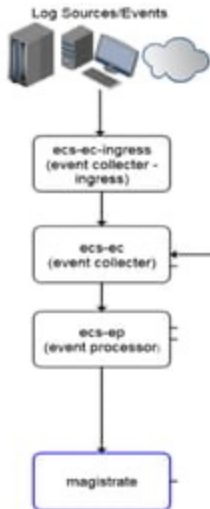
Host Profiling – Also called passive profiling or passive scanning. Watches flows on the network in order to make educated guesses about which IPs/assets exist and what ports are open.

Streaming – Responsible for the “real time (streaming)” view in User Interface

If an event matches a rule, the **Magistrate** component generates the response that is configure in the custom rule

Event  
Proces

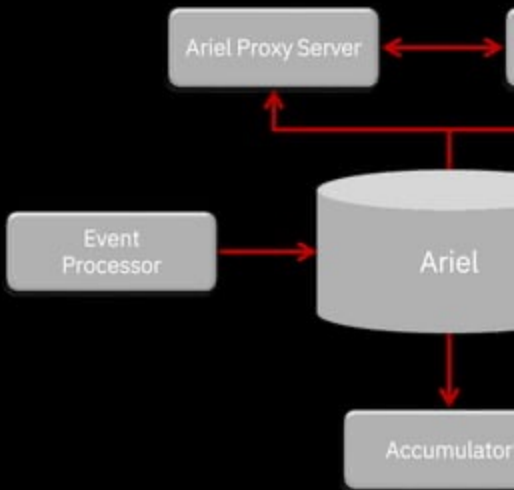
# QRadar Data Flow - Overall



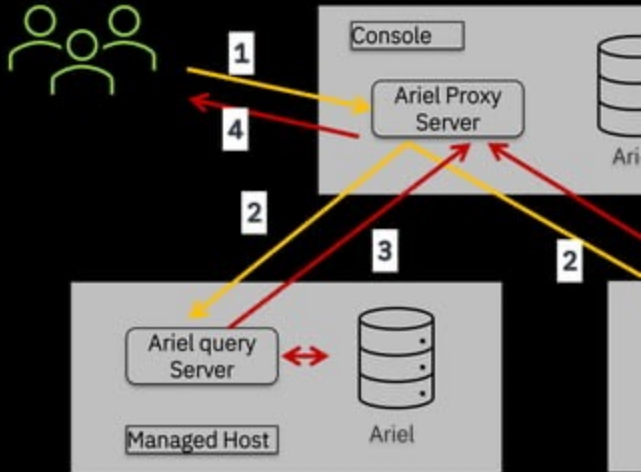
# Magistrate

- The Magistrate creates and stores **offenses** in the database. Offenses are then brought to the analyst's attention.
- The Magistrate instructs the **Ariel Proxy Server** to capture events and flows that triggered the creation of an offense.
- The Vulnerability Information Server (VIS) creates reports to existing assets based on information from the database.
- The Anomaly Detection Engine (ADE) searches for anomalies, which are then used for offense events.

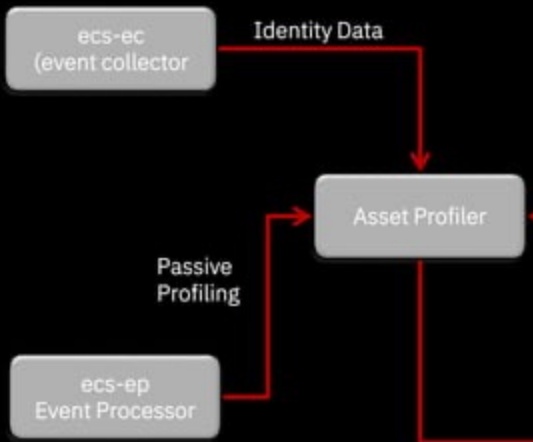
# Ariel Components



# Ariel Components



# Asset and Vulnerability Flow



# Gathering asset information

## Active scanners

QRadar Vulnerability Manager scanner, Nessus, Nmap, Qualys, and others

## Provide:

- List of hosts with risks and potential vulnerabilities
- IP and MAC addresses
- Open ports
- Services and versions
- Operating system

## Pros

- Detailed host information
- Policy and compliance information

## Cons

- Out of date quickly
- Full network scans can take weeks
- Active scanners cannot scan past firewalls
- User can hide from active scans

A  
S  
S  
E  
T  
  
P  
r  
O  
f  
i  
l  
e  
s

# The Remainder

Hostcontext

"Owns" the host it is responsible for  
processes and for overall system health

Reporting Executor

A stopwatch responsible for monitoring  
when they should run and for coordinating the  
runner

Report Runner

The process that actually runs the reports  
postgres, Ariel, etc.

Tomcat

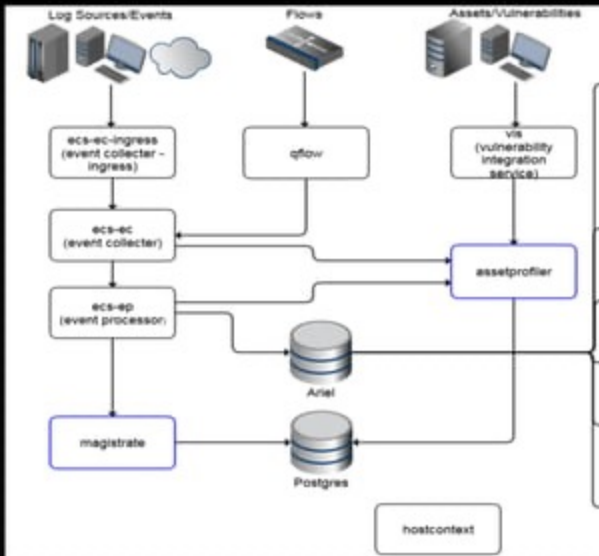
Process that drives the web interface

Historical Correlation  
Processor

Process that is responsible for running the  
specified search, runs on a schedule or  
on QRadar time or demand



# QRadar Data Flow - Overall



# Thank you

Follow us on:

[ibm.com/security](https://ibm.com/security)

[securityintelligence.com](https://securityintelligence.com)

[ibm.com/security/community](https://ibm.com/security/community)

[xforce.ibmcloud.com](https://xforce.ibmcloud.com)

[@ibmsecurity](https://twitter.com/ibmsecurity)

[youtube.com/ibmsecurity](https://youtube.com/ibmsecurity)

© Copyright  
information  
function re  
the IBM log  
the United  
of others.

Statement  
Detection  
information  
including fo  
product, se  
products  
additional  
warrant the  
or illegal co



**IBM Security**

IBM