Micro Focus Security
ArcSight Investigate

Software Version: 3.1.0

User's Guide

Document Release Date: April, 2020
Software Release Date: April, 2020

Legal Notices

Micro Focus
The Lawn
22-30 Old Bath Road
Newbury, Berkshire RG14 1QN
UK

https://www.microfocus.com

Copyright Notice

Confidential computer software. Valid license from Micro Focus required for possession, use or copying. The information contained herein is subject to change without notice.

The only warranties for Micro Focus products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Micro Focus shall not be liable for technical or editorial errors or omissions contained herein.

No portion of this product's documentation may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or information storage and retrieval systems, for any purpose other than the purchaser's internal use, without the express written permission of Micro Focus.

Notwithstanding anything to the contrary in your license agreement for Micro Focus ArcSight software, you may reverse engineer and modify certain open source components of the software in accordance with the license terms for those particular components. See below for the applicable terms.

U.S. Governmental Rights. For purposes of your license to Micro Focus ArcSight software, “commercial computer software” is defined at FAR 2.101. If acquired by or on behalf of a civilian agency, the U.S. Government acquires this commercial computer software and/or commercial computer software documentation and other technical data subject to the terms of the Agreement as specified in 48 C.F.R. 12.212 (Computer Software) and 12.211 (Technical Data) of the Federal Acquisition Regulation (“FAR”) and its successors. If acquired by or on behalf of any agency within the Department of Defense (“DOD”), the U.S. Government acquires this commercial computer software and/or commercial computer software documentation subject to the terms of the Agreement as specified in 48 C.F.R. 227.7202-3 of the DOD FAR Supplement (“DFARS”) and its successors. This U.S. Government Rights Section 18.11 is in lieu of, and supersedes, any other FAR, DFARS, or other clause or provision that addresses government rights in computer software or technical data.

Trademark Notices

Adobe™ is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Documentation Updates

The title page of this document contains the following identifying information:

- Software Version number
- Document Release Date, which changes each time the document is updated
- Software Release Date, which indicates the release date of this version of the software

To check for recent updates or to verify that you are using the most recent edition of a document, go to:

ArcSight Product Documentation on the Micro Focus Security Community

Support

Contact Information

Phone	A list of phone numbers is available on the Technical Support Page: https://softwaresupport.softwaregrp.com/support-contact-information
Support Web Site	https://softwaresupport.softwaregrp.com/
ArcSight Product Documentation	https://community.softwaregrp.com/t5/ArcSight-Product-Documentation/ct-p/productdocs

Contents

Chapter 1: Introduction 1

How ArcSight Investigate Works 1

ArcSight Investigate Features 3

Chapter 2: Searching Event Data 5

Specifying IP Address Ranges and Subnets as Query Input 9

Managing Search Results Fieldsets 10

Setting Default Fieldsets 11

Adding Data Visualizations 12

Creating Custom Visualizations 14

Managing Search Results Information 17

Adding a Lookup List to Extend Searches 19

Searching Events in ESM 21

About the Search Query Syntax 21

Query Syntax Requirements 22

Search Operators 24

Group Aliases 30

Field Aliases 30

Chapter 3: Data Quality Dashboard 36

Data Quality Aggregation 36

Data Quality Dashboard Categories 36

Data Quality Dashboard Visualizations 37

Chapter 4: Viewing Event Traffic in Host Profiler 39

Chapter 5: Managing the Dashboard 44

Chapter 6: Tracking DGA Activity with DNS Analysis 45

Configuring MS-DNS SmartConnector for DGA 45

Chapter 7: Auto Pass License 51

Installing Auto Pass License 52

Chapter 8: Analyzing Anomalous Data with Outlier Analytics 54

Defining and Building Models 54

Scoring a Model 56

Viewing Scored Data 58

Chapter 9: Integrating SOAR Applications with Investigate 61

Configuring Investigate for SOAR Integration 61

Invoking SOAR Actions from Investigate and Viewing Results 63

Chapter 10: Managing Users 66

Managing users 66

Creating a User 66

Importing Users from ESM 67

Managing User Groups 69

What Is a User Group? 69

Creating a User Group 70

Adding Users to a Group 70

Removing a User From a Group 71

Adding or Removing a Group Manager 72

Deleting a Group 72

Searching For a User 72

Managing a User's Account 73

Managing Roles 74

What is a Role? 74

Creating a Role 74

Editing a Role 75

Deleting a Role 76

Removing a Role From a User 76

Assigning a Role to a User 77

Built-in Roles 78

Permissions 78

Chapter 11: Processes 80

Terminating processes 80

Appendix A: FAQs 81

Appendix B: Debug Log Levels 83

Send Documentation Feedback 84

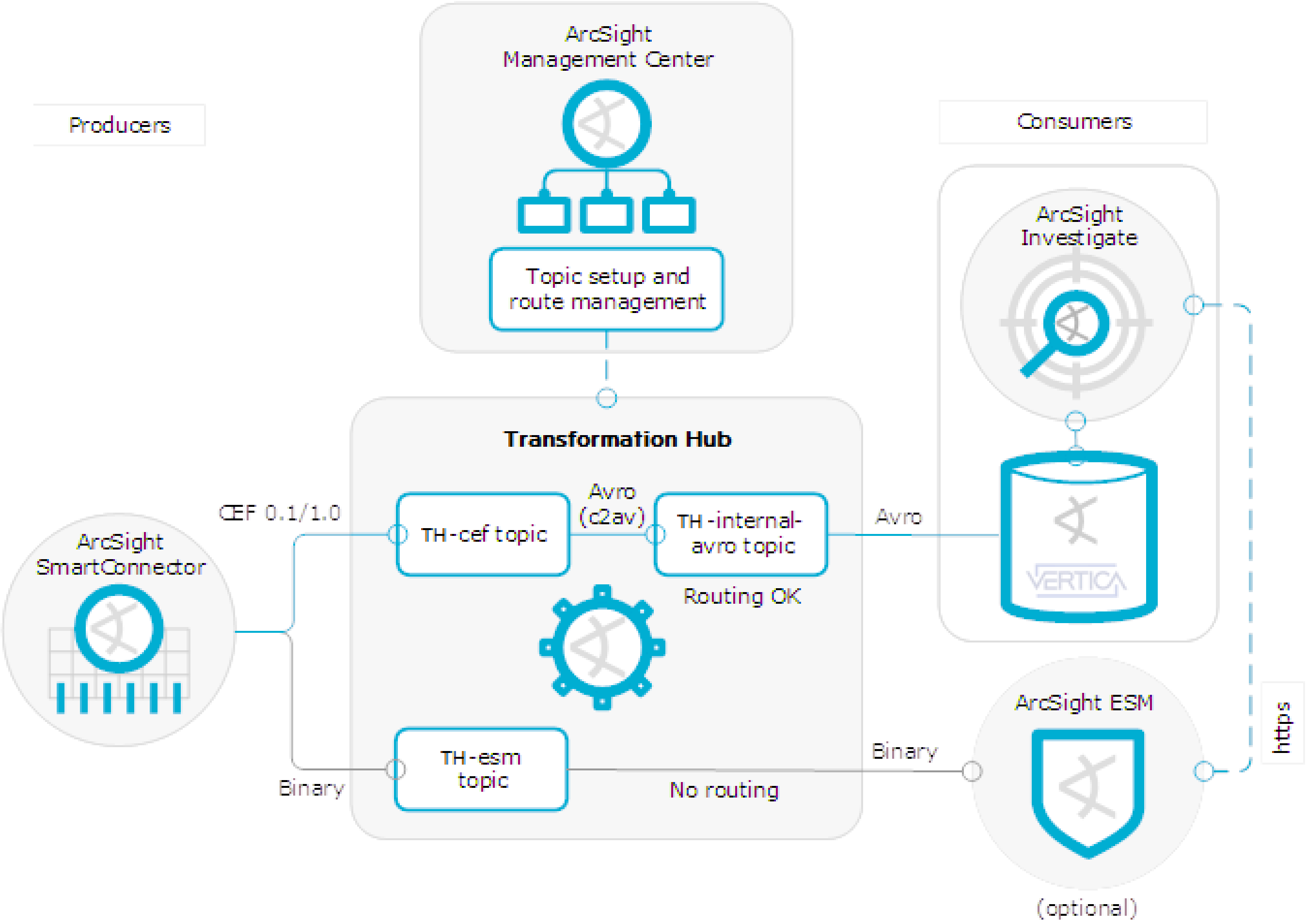
Glossary 85

Chapter 1: Introduction

ArcSight Investigate is a high-capacity data management and analysis engine that enables you to search, analyze, and visualize machine-generated data gathered from web sites, applications, sensors, and devices that comprise your monitored network. Investigate indexes the events from your data source so that you can view and search them. The intuitive search language makes it easy to formulate queries and then create reports and visualizations based on the search results. The information that a search yields can help you detect and investigate breaches before substantial damage occurs. From this, you can also evaluate the effectiveness of security policies and rules and security applications.

How ArcSight Investigate Works

The following image represents the Investigate architecture:



ArcSight Transformation Hub and ArcSight Smart Connectors are essential parts of the solution. Connectors send normalized and categorized common event format (CEF) events to the Transformation Hub topic `th-cef`. Transformation Hub transforms the events to Apache Avro format and then the Vertica Kafka scheduler consumes them and loads them into the Vertica database. reads the events from the Vertica database and then displays them on the **Search** page.

can extend the ArcSight Enterprise Security Manager (ESM) application to allow further investigation into events in an active channel. ESM generates a URL that opens with query input based on the data selected in the active channel.

The following components comprise the search function:

- **Search user interface**

  The **Search** page is where you start an investigation. It contains the **Search** bar, **Filter** field, **Timeline**, data visualization charts, and **Events** table.

- **Search backend**

  The search backend saves searches, user preferences, and proxy search requests to the search engine using the REST API.

- **Search engine**

  The search engine is a scalable server-side application that executes and caches large search queries in the Vertica database.

- **Vertica database**

  The database serves as the main data store, as well as a cache.

The Investigate web application includes the following pages:

- The **Dashboard** page allows you to view charts and includes text boxes for note-taking.
- The **Search** page allows you to search events and manage the search process.
- The **Insights** page allows you to view data for specified security use cases, such as Host Profiler.
- The **Configuration** page allows you to create and manage lookup lists, integrate SOAR applications, and build and score models that help you detect anomalous behavior.
- The **Integration** page allows you to view and manage SOAR notifications.
- The **Admin** page allows system administrators to create users and establish user rights.

ArcSight Investigate Features

Investigate provides the following features:

- ## Search

  Search is the primary way to navigate data in . The search is contextual and has auto-suggest capability to help you specify search criteria and improve productivity. The search function retrieves Vertica data rapidly. You can retrieve events from an index, use statistical commands to calculate metrics and generate charts, search for specific conditions within a rolling time window, identify patterns in your data, and predict future trends. You can generate charts in order to better understand the search results. Investigate supports up to 100 concurrent queries per installation, 10 active searches, and 40 saved searches per user. You can export a search either as a CSV or PDF file.

  For more information, see Searching Event Data.

- ## Indexing

  indexes machine data, including data streaming from packaged and custom applications, application servers, web servers, databases, networks, virtual machines, telecoms equipment, operating systems, and sensors that make up your IT infrastructure. The maximum indexing volume depends on the license.

- ## Data Analysis

  enables you to conduct a security investigation by filtering, comparing, visualizing, and analyzing event data dynamically. You can expedite the investigation process with quick and easy data analysis, deriving insights without any complexity. provides precise investigation outcomes through pre-defined queries (and fieldsets) for security use cases to improve SOC efficiency and reduce threat posture.

- ## Charting

  You can use the chart editor to graphically represent search results. The editor enables you to map attributes defined by data-model objects to a chart without having to write the searches to generate them.

  Investigate offers the following methods for building charts:

  - Built-in security analytics provide pre-defined charts that are configured for specific security use cases.
  - User-defined charts, where you can define all of the chart elements, including the type, fields, and functions used.

  Investigate saves charts with the search. You can also independently add charts to a dashboard.

  For more information, see Adding Data Visualizations, Creating Custom Visualizations, and Managing the Dashboard.

- **Dashboard**

  You can add charts and text boxes for taking notes to the dashboard.

  A chart can have a fixed start and end date, where you cannot refresh data, or a "canned" date range. For example, for a last-30-minutes "canned" date range, Investigate updates data based upon the most recent 30 minutes.

  For more information, see Managing the Dashboard.

- **Host Profiler**

  Host Profiler is a pre-defined dashboard where you can monitor event traffic for a specified host using visualization widgets. Investigate displays the traffic for the five most active host ports and communication paths related to other systems to help you to better analyze events.

  For more information, see Viewing Event Traffic in Host Profiler.

- **DNS Analytics**

  Pre-set visualizations enable you to monitor Domain Generating Algorithm (DGA) activity, often seen in malware.

  For more information, see Tracking DGA Activity with DNS Analysis.

- **Outlier Analytics**

  Outlier Analytics allows you to compare incoming EventCount, BytesIn, and BytesOut values to typical values for your environment in order to identify anomalous behavior.

  For more information, see Analyzing Anomalous Data with Outlier Analytics.

- **SOAR Integration**

  Investigate supports integration with selected SOAR (Security Orchestration, Automation, and Response) applications. Currently, you can integrate Investigate with Demisto, Operations Orchestration, or Siemplify Enterprise.

  For more information, see Integrating SOAR Applications with Investigate.

Chapter 2: Searching Event Data

The search function allows you to investigate specific alerts or search for events that meet certain criteria, and then view the results in tabular and graphical format so that you can detect anomalies that point to security threats.

When a search is run either from a search panel or any of the insight dashboards, a search status label and a search progress bar appear. The search progress bar has a pause (allows to pause a search) or a play button (allows to continue retrieving search results) next to it.

A search status label describes if the search is paused or running. It also tells how many results are retrieved so far. It also provides information from which time range the results are being read (if the search is running) or will be read (if the search is paused). If the number of search results exceeds 2M a warning icon will be shown in the status label.

A search progress bar shows the progress of search execution in terms of how many time chunks are left unprocessed. If the current chunk is the last one, the progress bar stays at 99 percent. If the number of results exceeds 2M, the progress bar turns yellow.

Investigate supports up to 10 active searches and up to 40 saved searches. Various filters allow you to refine your searches.

An event search consists of specifying query input, search result fields, and the time period for which you want to search events. Queries are case sensitive. The query input determines the search type (full text, natural language, or contextual). As you specify the criteria for a search query, Investigate suggests search items and operators based on a schema data dictionary. You can also choose from pre-defined queries.

A search query can either have a fixed start and end date, where you cannot refresh data, or a "canned" date range. For example, for the last 30 minutes "canned" search, Investigate updates data upon re-executing the search based on the most recent 30 minutes.

> **Note:** When performing a search with two or more identical queries (be it canned and / or explicit) the number of events returned for the second search will correspond to the next chunk. If the search is resumed, the first search will be moved to the next chunk as well, maintaining the same number of events retrieved.

If an event does not have data for a schema field, Investigate represents the absence of data (null) in various ways:

Search field	`Null`, `NULL` and `null` query formats are supported
Events table	Empty cell

Charts	*(Null)*
Empty field from ESM (`name=’’`)	`name = ‘’, NULL`

Refreshing the browser as you update a search does not save your changes. You must first click **Save** on the task bar. If you navigate away from the search to one of the following pages, Investigate automatically saves the search:

- Dashboard
- Insights
- Configuration
- Integration
- Another search

If you make changes to a search query, a fieldset, or the range selector, Investigate does not save your changes until you click **Apply**.

Investigate does not automatically save a search in the following situations:

- You close the browser or a tab.
- You log out.

For searches that you create in a different timezone, **Timeline** converts the time segments to local times. If a chart or the **Events** table includes a time attribute, Investigate converts the time to local time. However, the aggregation reflects the original timezone. For example, if **Timeline** has seven bars in the original timezone, the number of bars could increase or decrease to reflect the current timezone.

To create a new search:

1. In the left navigation, click **Search** > **New Search**.
2. Accept the default search name or rename the search.
3. In the search bar, specify the query input.

   For example, `Source Address = 192.10.11.12 and Destination Address less than 192.10.11.12`.

   To use a canned query, type # and select the query.

   Investigate treats a comma (,) between search items and values as an `or` operator.

   To search for a field without data, use the `null` field value.

   You can specify IPv4, IPv6, and MAC adresses. For more information, see Specifying IP Address Ranges and Subnets as Query Input.
4. Accept the default fieldset for the search results or click **Default Fieldset** to change the fieldset.

Note: Depending on your data access permissions, you might not see all of the possible fields. To view the fields for which you have access, from the left navigation pane select <User_Name> **My Profile** > **Data Access**.

For information about changing fieldsets, see Managing Search Results Fieldsets.

5. Accept the default time range (**Last 30 minutes**) or click the time drop-down list and use the **Custom Range** fields to specify a different time range, and then click **Search**.

   > **Note:** Searches for events in a time range are based on the timestamps of matching events and use the time zone of the local browser. You might need to account for the time zone offset from UTC and from other time zones, including Daylight Savings Time.
   >
   > The time range that you specify in the time range selector is inclusive. Investigate includes the whole second as the end time. For example, if you specify a time range between 2018-01-01 12:00:00 and 2018-01-01 12:59:59, Investigate includes all data from 2018-01-01 12:00:00.000 to 2018-01-01 12:59:59.999, inclusive.

   Investigate populates the **Timeline** and **Events** table.

   Depending on the number of events that Investigate retrieves, the search might pause to indicate that the amount of data might impact the search performance. You might want to select a smaller time range. To resume a search, click the play button in the progress bar.

   To cancel a search, click the pause button in the progress bar.

6. If you want to save the search for future use, click **Save**.

   Searches that you save are available from the left navigation pane.

For information about changing the layout of the **Events** table and viewing data details in the table, see Managing Search Results Information.

Specifying IP Address Ranges and Subnets as Query Input

Investigate stores IPv4, IPv6, and MAC addresses in a format that provides more flexibility and better search performance. This format allows you to:

- Compare IP addresses with optimum performance. For example:

  `Agent Address > 192.10.11.12`
- Specify a range of IP addresses. For example:

  `Agent Address in between 192.2.13.1 and 192.2.13.11`

  `Source Address greater equal than 192.10.11.12 and Destination Address less than 192.112.98.33`
- Use abbreviated input search notation. For example:
  - `Agent Address in subnet 192.*`
    For this IPv4 address, the input specifies IP addresses in the subnet starting with 192.
  - `Agent Address in subnet 192.0.0.0/8`
    For this IPv4 address, the input specifies an agent address in a subnet that uses CIDR notation. The first eight bits are the network part of the address, leaving the last 24 bits for specific host addresses.
  - `Agent Address in subnet 2001:0db8:0000:0000:0000:ff00:0042:8329/24`
    For this IPv6 address, the input specifies an agent address in a subnet that uses CIDR notation. The first 24 bits are the network part of the address, leaving the last 40 bits for specific host addresses.

Investigate supports the following regular formats:

- `aa:aa:aa:aa:aa:aa`
- `aa-aa-aa-aa-aa-aa`
- `aaaa.aaaa.aaaa`

Investigate supports MAC addresses in IPv6 EUI-64 format (see RFC 2373):

- `fe80:0000:0000:0000:aaaa:aaff:feaa:aaaa`
- `FE80::aaaa:aaff:feaa:aaaa`

When Investigate stores MAC addresses, it converts them to IPv6 format. For example, `B9:0D:78:10:40:DA` becomes `fe80:0000:0000:0000:bb0D:78ff:fe10:40DA`.

Investigate supports IPv4 addresses in `a.b.c.d` format. To specify an IPv4 address in a subnet, use `a.*`, `a.b.*`, `a.b.c.*`, or `a.b.c.d/8`.

Investigate supports IPv6 addresses in full form and canonical form (see RFC 5952). For example:

- Full form: `2001:0db8:0000:0000:0000:ff00:0042:8329`
- Canonical form without leading zeroes in each group: `2001:db8:0:0:0:ff00:42:8329`
- Canonical form without consecutive sections of zeroes: `2001:db8::ff00:42:8329`

To specify an IPv6 address in a subnet, use any of the formats above with CIDR notation. For example:

```
2001:0db8:0000:0000:0000:ff00:0042:8329
```

```
2001:0db8:0000:0000:0000:ff00:0042:8329/24
```

```
2001:db8::/32
```

For `2001:db8::/32`, you can omit part of the IPv6 address, depending on the subnet that you are querying.

Managing Search Results Fieldsets

The fieldset determines the search result fields that are visible in the **Events** table and available for creating visualizations. The default fieldset contains the most common event fields, but additional fields are available. Each field can provide the 10 most and least common values. Multiple searches can share a fieldset. You can customize the default fieldset for individual searches, and you can add lookup list fields to a fieldset.

To create a fieldset:

1. On the **Search** page, in the search bar, click **Base Events Fields**.
2. From the drop-down menu, select **Create a new set**.
3. Select and/or deselect the desired fields.

   To view the complete list of available fields, click **View all**.

   To locate a specific field, use the search field.
4. To add lookup list fields to the fieldset, click **Lookup Lists**.
5. Accept the default name for the new fieldset or specify a name, and then click **Save**.

To edit a fieldset:

1. In the search bar, click the fieldset name.

   > **Note:** The fieldset name is the name of the last used fieldset.
2. If the last used fieldset is not the fieldset that you want to edit, select another fieldset from the drop-down menu.
3. From the drop-down menu, select **Edit this set**.
4. Select and/or deselect the desired fields.

> **Note:** When you remove a field from a fieldset, Investigate removes all filters and charts that use that field.

5. Make any other desired changes, such as adding lookup list fields or renaming the fieldset, and then click **Save**.

To delete a fieldset:

> **Note:** Fieldsets can be deleted as long as they are user created, and they haven’t been designated as the default fieldset. If this happens the fieldset will be displayed as ***Custom*** for the remaining active searches.

1. In the search bar, select the fieldset that you want to delete.

   > **Note:** The fieldset name is the name of the last used fieldset.

2. If the last used fieldset is not the fieldset that you want to edit, select another fieldset from the drop-down menu.
3. From the drop-down menu, click **Edit this set**.
4. Click **Delete**.

Setting Default Fieldsets

Setting a default fieldset will improve search performance (search results will display faster) by retrieving less fields. Minimizing the number of fields in the default fieldset will not compromise the required fields.

Requirements

- Select a new fieldset other than the default **Base Event Fields.**
- The Admin user is the only one who can set the default fieldset for all Investigate users.
- Only one fieldset can be designated as the default fieldset. There must be a default fieldset
- Saved fieldsets are the only ones that can be set as default.
- Each fieldset should have a unique name (There can’t be 2 fieldsets with the same name) and it is not case sensitive.
- A default fieldset cannot be edited and saved under the original name

> **Note:** The **Default Fieldset**, and the **Base Event Fields** have been changed to **Custom Base Event Fields** after the Investigate upgrade from 2.40 to 3.0.

The following fields have changed in the fieldset:

Removed Fields	Added Fields
Agent Receipt Time	Base Event Count
Destination User Privileges	Destination Hostname
Source User Privileges	Bytes Out
Destination User ID	Bytes In
Source User ID	Category Behavior
Agent Hostname	
Agent ID	

Adding Data Visualizations

To better understand search results data, you can represent it graphically on the **Search** page. Investigate allows you to add up to 10 data visualizations.

Investigate provides data comparison visualizations and non-comparison visualizations. Data comparison visualizations include line, column, bar, and area charts. Non-comparison visualizations include pie and scatter plot charts.

> **Note:** To display the tooltip text, hover over the Y-AXIS values.

You can add **Search** page visualizations to the dashboard as widgets. For more information, see Managing the Dashboard.

Investigate provides the following pre-defined visualizations:

Authentication Activity
Login by Destination Address Over Time
Login by Destination Username Over Time
Login by Username
Login Over Time
Source Activity
Bytes Out by Source Address
Destination Hostname by Source Address - Detailed
Destination Hostname by Source Address - Summary
Destination Port by Source Address - Detailed
Destination Port by Source Address - Summary

Source Antivirus Activity
Top Source Addresses
Destination Activity
Bytes In by Destination Address
Bytes Out by Destination Address
Bytes Out by Destination Hostname
Bytes Out by Request URL
Destination Antivirus Activity
Destination Port by Destination Address
Source Address by Destination Address - Detailed
Source Address by Destination Address - Summary
Top Destination Addresses
Top Destination Hostname
Port & Protocol Activity
Bytes In by Destination Port
Bytes Out by Destination Port
Secure Communication Ports-Bytes Out by Destination Hostname
Secure Communication Ports-Bytes Out By Source Address
Top Destination Ports
General
Authorization Changes by Destination Address
Bytes In by Destination Username
Bytes In Over Time
Bytes Out by Destination Host and Source Username
Bytes Out by Device Vendor
Bytes Out by Source Username
Bytes Out Over Time
Events Count Over Time
Top Device Vendors
DNS Activity
DNS Analysis: Top Hosts
Top Hosts by DNS Events Sum Bytes Out

Top Hosts by Number of Unique DGA Domains		
Top DGA Domains by Number of Unique Hosts		
DNS Analysis Over Time		

To add a pre-defined visualization to the Search page:

1. Expand the **Visualize** area, and then click **Create Visualizations**.
2. Select the desired category, and then select the desired visualizations.

For information about creating custom visualizations, see Creating Custom Visualizations.

Creating Custom Visualizations

Line, bar, column, and area charts are data comparison visualizations. For these visualizations, you can create up to six series of data comparisons.

The first chart series sets the X- and Y-axis parameters, which remain set for any subsequent series. Ordering that you apply to the first chart series applies to subsequent series. For subsequent series, you can filter by different fields and set aggregate functions for the X and Y axis parameters.

The X axis can receive fields with a continuous value. Investigate applies the `sum ()` aggregate function to continuous-value fields, and converts discrete-value fields to continuous value by applying the **`count()`** aggregate function. The Y axis can receive multiple discrete fields. applies the **`count()`** aggregate function to continuous-value fields. You can change the aggregate function.

Non-comparison visualizations include pie and scatter plot charts.

ArcSight provides the following X and Y axis options (for bar charts, the X and Y axis behavior is reversed):

Field Type	X Axis Function	Y Axis Function
Time	second minute (default) day hour week month year value itself	count count distinct For example, count the number of events for the time period.
String	value itself	count (default) count distinct value itself (scatter chart and bar chart)
Number	value itself	count count distinct sum (default) average max min Number value itself (only for scatter plot) **Note:** For the average function, the default is the arithmetic mean. For example, for bytes out, the average will be sum (BytesOut) / number of events that contain BytesOut. If you select **Group by User**, Investigate uses the formula sum (Bytes Out (only for events when `user!=Null`) / distinct number of users (without `Null`).

When you drag a discrete-value field to a continuous-value parameter, converts the field to a continuous-value field. For example, for **File Name**, Investigate applies the `count()` function.

Within a parameter, Investigate displays fields in the following formats:

- Single key/value pair: `<field>:<value>`

  For example, `department:sales`.
- Single key with multiple values: `<field>:<value1>,<value2>,...`

  For example, `user:johnny, bob,...`.
- Aggregate function: `<function>(<field>)`

  For example, `- sum(Bytes Out)` or `- month(Event Time)`.

To create a line, bar, column, or area chart:

1. On the **Search** page, expand the **Visualize** area, and then click **Create Visualizations**.
2. Click **Create New**, and then select the desired chart type.

   The available fields depend on the fieldset that is currently in use. To change the available fields, click the fieldset name, and then select the desired fieldset.
3. From the list of available fields, drag the desired fields to **X-Axis** and **Y-Axis**.
4. To compare event field data against the entire dataset, drag the desired field to **Filter By**.

   The parameter can receive multiple discrete fields. By default, Investigate applies all values for a field. To change the field values, click the field and specify the desired field values.
5. To specify the field by which to sort records, click **Order By**.

   The sort order is dependent on the **Y-Axis** field. By default, Investigate displays records in ascending order.
6. For a horizontal bar visualization, specify segmenting of Y axis bars by dragging the desired field to **Category**.

   Categories allow you to specify a secondary discrete-value field. Investigate segments each bar in the Y axis by the secondary category.
7. To set a baseline by which to compare data, select **Plot Line**, and then specify the baseline value in the adjacent field.
8. To create another data segment comparison, click **Add Series** and specify any new parameters and aggregate functions.

   If you change **Order By** for any series, Investigate updates all chart series.
9. When you are ready to create the visualization, click **Done**.

To create a pie chart:

1. On the **Search** page, expand the **Visualize** area, and then click **Create Visualizations**.
2. Click **Create New**, and then select **Pie Chart**.

   The available fields depend on the fieldset that is currently in use. To change the available fields, click the fieldset name, and then select the desired fieldset.
3. From the list of available fields, drag the field by which you want to measure to **Measure**.

   Measure determines the size of the pie slices. The parameter can receive a continuous-value field. Investigate applies the aggregate function that is supported for the field.
4. From the list of available fields, drag the field that you want to use as the label to **Label**.

   The parameter can receive a discrete-value field. Investigate groups events by unique values for the field that you select.
5. To filter the chart, drag the field by which you want to filter to **Filter By**.

   The parameter can receive multiple discrete-value fields.
6. When you are ready to create the visualization, click **Done**.

To create a scatter plot chart:

1. On the **Search** page, expand the **Visualize** area, and then click **Create Visualizations**.
2. Click **Create New**, and then select **Scatter Plot**.

   The available fields depend on the fieldset that is currently in use. To change the available fields, click the fieldset name, and then select the desired fieldset.
3. From the list of available fields, drag the desired fields to **X-Axis** and **Y-Axis**.
4. From the list of available fields, drag the desired field to **Category**.

   In the scatter plot, a different color point represents each unique value for that field.
5. When you are ready to create the visualization, click **Done**.

To change the aggregate function for a parameter:

1. Click the field on the X or Y axis.
2. Select the desired value type from the **Aggregate values using** drop-down list.

After you save a visualization, click **...** to rename, edit, delete, or add the visualization to the dashboard. If you delete a visualization that you added to the dashboard, the visualization remains in the dashboard.

Managing Search Results Information

Investigate provides the following options for changing the layout of the **Events** table:

- Pin columns.

  To pin a column in order to better compare the column values against those of other columns, right click the column heading and select **Pin Column**. Investigate moves the column to the extreme left. You can pin multiple columns. To unpin a column, right-click the column heading and select **Unpin Column**.

- Remove columns.

  To remove columns from the **Events** table, click the wrench icon, and then deselect the columns that you do not want to view. You can also right-click a column heading and select **Hide Column**.

- Reorder columns by dragging them to new positions.
- Sort columns in ascending or descending order.

Investigate provides the following options for viewing data details:

- View the most and least common values for an event record field.

  To help filter data for security threats, you can quickly display the most and least common values for a field. When you right-click a column heading and select **Preview Top/Bottom**, Investigate displays the count and percentage of hits for the value. For example, the **Device Vendor** field might have a top value of "bluecoat" with a count of 3,000 hits, accounting for 30 percent of 10,000 results.

- View authenticated users.

  To view users who have successfully authenticated to an IP address or host name in the last 24 hours, right-click an IP address or host name and select **Get Authenticated Users**.

  Note: The fieldset for the original search must include **Device Receipt Time**.

- View all record fields for an event.

  You can expand each row in the **Events** table to view details for the event. Within the event details, you can expand the fields for more information.

- View all event data for a field value.

  To view all of the event data that is based on a field value, right-click a value in a table row, and then select **Search for**.

You can export search results to a PDF or CSV file. For PDF files, Investigate exports details about the search conditions and also exports visualizations. For CSV files, Investigate only exports **Events** table data. For both types of files, Investigate exports data based on the fieldset for the search.

To export to a PDF file, click **Export to PDF** at the top of the **Search** page. To export to a CSV file, click the CSV icon in the **Events** area.

Adding a Lookup List to Extend Searches

The lookup list feature allows you to create additional tables with different fields and store them in the Vertica database. You can add lookup list fields to fieldsets and use them in search queries.

Lookup lists must be CSV files that meet the following requirements:

- The first row must be a comma-separated list of field names. The field names cannot exceed 40 characters. They can only contain alphanumeric characters and underscores and must start with an alpha character.
- The remaining rows must be comma-separated values for the fields in the first row.
- All rows must contain the same number of values.
- You must select one of the columns as the key field, and the values of the key field must be unique.

  The key field is the field that you can use with the `in list` operator in queries.
- The file cannot exceed 25 fields and 2 million rows.
- The file cannot exceed 150 MB.

To create a lookup list:

1. In the left navigation pane, select **Configuration** > **Lookup Lists**, and then click **Add**.
2. Drag and drop your CSV file to the **Lookup Lists** page or click **Browse** to navigate to it.
3. Specify a lookup list name.

   The name cannot exceed 20 characters and can only contain alphanumeric characters and underscores. The name must start with an alpha character.

   Investigate displays all the fields of the lookup list, each with the default value type of `text`.
4. Specify the key field and either accept the value type or specify a different one.

   The following are possible values:

domain	
float	For a number whose radix point can be placed anywhere relative to the significant digits of the number
hostname	Fully qualified domain name
int	Integer value
ipv4	IPv4 address
ipv6	IPv6 address
mac	MAC address
short text	Text that cannot exceed 1K of space
long text	Text that cannot exceed 4K of space
time	Timestamp
url	Cannot exceed 4K
username	A string type

5. To upload the file as a table in the Vertica database, click **Upload**.

To replace a lookup list:

> **Note:** Replacing the contents of a lookup list does not affect queries that use the original lookup list. You cannot change the lookup list name. The field names in the replacement file must match the field names in the original file.

1. In the left navigation pane, select **Configuration** > **Lookup Lists**.
2. Select the desired list, and then click **Replace**.
3. Select the CSV file that you want to use to replace the contents of the existing lookup list.

To delete a lookup list:

1. In the left navigation pane, select **Configuration** > **Lookup Lists**.
2. Select the desired list, and then click the trash can icon.

Searching Events in ESM

From ArcSight Enterprise Security Manager (ESM), you can initiate a search in Investigate for a maximum of five fields. Within Investigate, you can filter ESM data for more specific results. You must enable Investigate in ESM. For more information, see the ESM Installation Guide.

To generate a search of Investigate events from ESM:

1. In ESM, open an active channel or view event details in the Inspect/Edit panel.
2. Right-click a row and make the appropriate selections.

   > **Note:** ESM fields that are not supported in searches are disabled.

3. To create a search from a specific value, select **ArcSight Investigate**.
4. To create a search from multiple values, select **ArcSight Investigate (Multiple Fields)**.

   You can select up to five fields, based on the available columns on the active channel.

ESM generates a URL, opens a browser, and creates a new search in .

About the Search Query Syntax

To perform a search

Location: Search page > search field

1. Enter the desired query criteria (keywords or information for which you are searching).
2. Select the desired time range.

3. Click **Search**.

   ArcSight searches for data that matches the criteria that you specified and displays the results.

For details, see the ***Micro Focus Security ArcSight Investigate User's Guide***

Type	**Description**	**Syntax**
Full Text Search	Searches across all columns using a 'contains' operation to determine if the value is found.	<value> Example: `ssh`
Field-Based Search	Searches based on the field and operator designation to determine if the value is found in the specified field.	<key> <operator> <value> Example: `sourceAddress = 10.0.111.5`
Hashtag Search (preset searches)	These are predefined queries that are referenced in the search input field using a hashtag, plus the name. In addition to predefined searches, the session searches and save searches can be used in the input field using a hashtag prefix.	

Query Syntax Requirements

Behavior	**Full text search**	**Field-based search**	**Hashtag search**
Case sensitivity	Case-insensitive	Case-insensitive	Case-insensitive
Exact Match	Keyword treated as keyword*. Example: `/Execute matches: /Execute, /Execute/Start, /Execute/Response,/Execute/Query`	Enclose value in double quotes. Example: `Category Behavior =“/Execute”`	N/A
Nesting, including parenthetical clauses, such as (a OR b) AND c	Allowed. Use Boolean operators to connect and nest keywords.	Allowed. Use Boolean operators to connect and nest keywords.	Allowed. Use Boolean operators to connect and nest keywords.

Behavior	Full text search	Field-based search	Hashtag search

Implicit Operators	When two values entered separated by a space, this is treated as an implicit AND condition.
	Example: `ssh fail`
		The AND/OR treatment depends on the operator used in the search.
		For example:
		`destinationAddress = 1.1.1.1, 2.2.2.2`
		is equivalent to
		`destinationAddress = 1.1.1.1 or destinationAddress = 2.2.2.2`,
		while the following query:
		`destinationAddress != 1.1.1.1, 2.2.2.2`
		is equivalent to:
		`destinationAddress != 1.1.1.1 and destinationAddress != 2.2.2.2`
			N/A

List Operations	N/A	Performs an inner join or a left join against a custom list.
		Syntax for Inner Join:
		`source address in list CustomListName_ CustomColumn Name`
		Syntax for Left Join:
		`source address not in list CustomListName_ CustomColumnName`
			N/A

Time format, when searching for events that occurred at a particular time	No specific format. The query needs to contain the exact timestamp string. For example, "10:34:35".	Use this format to specify a timestamp in a query:
		`YYYY-MM-DD`
		`YYYY-MM-DD HH:mm`
		`YYYY-MM-DD HH:mm:ss.fff`
		Use the in between `><`, greater than (`>`) or less than (`<`) operators to narrow the time range.
			N/A

Search Operators

The following table describes the operators you can specify in the **Search** field.

Operator	Operator Alternatives	Example
AND		#Firewall drop and sourceAddress equals 10.0.112.9 sourceAddress equals 10.0.112.9 and destinationAddress = 10.0.116.148
OR		fail OR ssh destinationAddress = 10.0.111.5 OR destinationAddress=10.0.116.148 destinationAddress = 10.0.111.5, 10.0.116.48
not equal	<> !=	destinationPort not equal 21
equals	= == is equal to equal	name equals INVALID password device vendor equals CISCO
greater than	> is greater	bytes In greater than 100
less than	< is less is lower less	bytes out less than 1000
greater equal than	>= gte greater equal	End Time greater equal than 2017-07-25 End Time greater equal than 2017-07-25 09:07 End Time greater equal than 2017-07-25 09:07:43 End Time greater equal than 2017-07-25 09:31:22.685
less equal than	<= lte less equal	Base Event Count less equal than or equal 50
starts with	startswith	message starts with FIN
does not start with		name does not start with FIN
ends with	endswith	message ends with out
does not end with		message does not end with out

Operator	Operator Alternatives	Example
contains	contain like has substring	name contains TCP
does not contain	does not have	name does not contain TCP
in list	match in list of	device vendor equals CISCO and source address in list customListName_customColumnName device vendor equals CISCO and source address in list badGuyIpList_badGuyIp
not in list	not match not in list of	source address not in list customListName_customColumnName source address not in list badGuyIpList_badGuyIp
in subnet	N/A	source address in subnet 10.0.0.0/8
not in subnet	N/A	source address not in subnet 10.0.0.0/8
| (Pipeline operator)	N/A	ssh | eval test1 = abs (40) ssh | eval test1 = sin (Bytes In)
eval <expression> name	N/A	| eval URL_Length = length (Request URL)
rename	N/A	| rename source address as src

Functions supported for eval operations

Function	Description	Example
abs(X)	Takes a number, X, and returns its absolute value. X can be a number, field or expression	The function assigns the evaluated value to the new field. If the value of X is 3 or -3, the function assigns the evaluated value of 3 to the field absnum. ... | eval absnum=abs(number) ...| eval absnum = abs(bytesIn) ...| eval absnum = abs(1 - bytesIn)
ceiling(X)	Rounds a number, X, up to the next highest integer X can be a number, field or expression	The following example returns n=2. ... | eval n=ceil(1.9)
exp(X)	Takes a number, X, and returns e^X. X can be a number, field or expression	The following example returns y=20.0855369231877. ... | eval y=exp(3)
floor(X)	Rounds a number, X, down to the nearest whole integer. X can be a number, field or expression	The following example returns 1. ... | eval n=floor(1.9)

Function	Description	Example
length(X)	Returns the character length of a string, X.	The following example returns the length of(field). If the field is 256 characters long, it returns n=256 ... | eval n=length(field) The following example returns n=3. (abc is a literal string,surrounded by double quotes.) ... | eval n=length("abc")
ln(X)	Takes a number, X, and returns its natural log. X can be a number, field or expression	The following example returns the natural log of the value of "bytesIn". If "bytesIn" contains 100, it returns 4.605170186. ... | eval lnBytes=ln(bytesIn)
log10(X)	Evaluates the log of number X with base 10. X can be a number, field or expression	The following example returns 4. ... | eval num=log10(10000).
log(X, Y)	Returns the logarithm to the specified base of the argument. X is the base and Y can be a number, field or expression X is optional. If not specified, it will take 10 as the default value.	The following example returns 0.301 ... | eval test1= log (10,2) The following example returns 0.301 as it takes the default base as 10 ... | eval test1 = log (2)
lower(X)	Takes a string argument, X, and returns the lowercase version.	The following example returns the value of the field username in lowercase. If the username field contains FRED BROWN, it returns name=fredbrown. ... | eval name=lower("username")
mod(X, Y)	Returns the modulo of X and Y. (X%Y; the remainder of X divided by Y.)	The following example returns 5. ... | eval newField = mod(25,10)
randomint (X)	Returns a random number between 0 and X-1 X can be any positive integer between the values 1 and 9,223,372,036,854,775,807.	The following example will return a random number between 0 and 9. ... | eval newField = randomint(10)
round(X, Y)	Rounds X to the nearest integer. Y is the precision to use, if omitted the default precision is zero. X can be a number, field or expression Y is a numeric value to indicate the precision.	The following example returns 1. ... | eval n=round(1.4) The following example returns 2. ... | eval n=round(1.5)
sqrt(X)	Takes one numeric argument, X, and returns its square root. X can be a number, field or expression	The following example returns 3. ... | eval n=sqrt(9)

Function	Description	Example
substr (X,Y,Z)	This function returns a new string that is a substring of string X. The substring begins with the character at index Y and extends up to the character at index Z-1. The index is a number that indicates the location of the characters in string X, from left to right, starting with zero. Y - can be negative. Z - cannot be negative.	The following example returns"g". ...| eval n=substr("ArcSight",5,6) The following example returns"cSig". ...| eval n=substr("ArcSight",2,6) The following example returns"Arc". ...| eval n=substr("ArcSight",0,3)

trim(X)
ltrim(X)
rtrim(X)
	trim(X) removes all spaces from both sides of the string X.
	ltrim(X) removes all spaces from the left side of the string X.
	rtrim(X) removes all spaces from the right side of the string X.
		For the sake of the example, assume that X is a literal string and _ represents any number of space characters.
		The following example returns trimmed= "string_".
		... | eval trimmed=ltrim("_string_")
		The following example returns trimmed="_string".
		... | eval trimmed=rtrim("_string_")
		The following example returns "string".
		... | eval trimmed=trim("_string_")

upper(X)	Takes one string argument and returns the uppercase version.	The following example returns the value of the field username in uppercase. If username contains fred brown, it returns name=FRED BROWN. ... | eval name=upper("username")
cos(X)	Takes one numeric argument, X, and returns its trigonometric cosine	The following example returns 2435538 ...| eval newField = cos(3)
sin(X)	Takes one numeric argument, X, and returns its trigonometric sine	The following example returns 0.141120008059867 ...| eval newField = sin(3)
cot(X)	Takes one numeric argument, X, and returns its trigonometric cotangent	The following example returns -7.01525255143453 ...| eval newField = cot(3)
tan(X)	Takes one numeric argument, X, and returns its trigonometric tangent	The following example returns -0.142546543074278 ...| eval newField = tan(3)

Function	Description	Example
tanh(X)	Takes one numeric argument, X, and returns its hyperbolic tangent	The following example returns 0.99505475368673 ...| eval newField = tanh(3)
sinh(X)	Takes one numeric argument, X, and returns its hyperbolic sine	The following example returns 10.0178749274099 ...| eval newField = sinh(3)
cosh(X)	Takes one numeric argument, X, and returns its hyperbolic cosine	The following example returns 10.0676619957778 ...| eval newField = cosh(3)
acos(X)	Takes one numeric argument, X, and returns its trigonometric inverse cosine	The following example returns 1.2661036727795 ...| eval newField = acos(0.3)
asin(X)	Takes one numeric argument, X, and returns its trigonometric inverse sine	The following example returns 0.304692654015398 ...| eval newField = asin(3)
atan(X)	Takes one numeric argument, X, and returns its trigonometric inverse tangent	The following example returns 0.291456794477867 ...| eval newField = atan(3)
atan2(X, Y)	Returns a value representing the trigonometric inverse tangent of the arithmetic dividend of the arguments.	The following example returns 1.10714871 ...| eval newField = atan2(2,1)
cbrt(X)	Takes one numeric argument, X, and returns its cube root	The following example returns 8 ... | eval n=cbrt(2)
sign(X)	Returns a value of -1, 0, or 1 representing the arithmetic sign of the argument.	The following example returns -1 ... | eval newField = sign(-8.4) The following example returns 1 ... | eval newField = sign(4) The following example returns 0 ... | eval newField = sign(0)

Function	Description	Example
trunc(X, Y)	Returns the expression value truncated (toward zero) X can be a number, field or expression Y is a numeric value to indicate the precision.	The following example returns 1 ... | eval newField = trunc(1.9) The following example returns 2.89 ... | eval newField = trunc(2.89999, 2)
greatest (X, Y [, Z, N, ...])	Returns the largest value in a list of expressions. The list is up to 20 elements long. In the list of expressions all elements must be of same type. The only supported types are numeric and string. X can be a number, field or expression.	The following example returns 9 ... | eval newField = greatest(7, 5, 9) The following example returns site ... | eval newField = greatest('sit', 'site', 'sight') The following example returns 100 when bytesIn is less than 100 ... | eval newField = greatest(bytesIn, 100)
least(X, Y [, Z, N, ...])	Returns the smallest value in a list of expressions. The list is up to 20 elements long. In the list of expressions all elements must be of same type. The only supported types are numeric and string. X can be a number, field or expression.	The following example returns 5 ... | eval newField = least(7, 5, 9) The following example return sight ... | eval newField = least('sit', 'site', 'sight') The following example returns 100 when bytesIn is greater than 100 ... | eval newField = least(bytesIn, 100)
coalesce (X[, Y, Z, N, ...])	Returns the value of the first non-null expression in the list. If all expressions evaluate to null, then COALESCE returns null. The list is up to 20 elements long. In the list of expressions all elements must be of same type. The only supported types are numeric and string. X can be a number, field or expression	The following example returns 2 ... | eval newField = coalesce(null, null,2,3)
power(X, Y)	Returns a value representing one number raised to the power of another number. X is the base and Y the exponent. X and Y can be a number, field or expression	The following example returns 8 ... | eval newField = power(2, 3)

Function	Description	Example
md5(X)	Calculates the MD5 hash of string, returning the result as a VARCHAR string in hexadecimal. X must be a string	The following example return 202cb962ac59075b964b07152d234b70 ... | eval newField = md5('123')
nullif(X,Y)	Compares two expressions. If the expressions are not equal, the function returns the first expression (expression1). If the expressions are equal, the function returns null. X and Y can be a number, field or expression. Y must have same data type that X.	The following example returns 2 ... | eval newField = nullif(2, 3) The following example returns null ... | eval newField = nullif(2, 2)
isnull(X)	Returns true if the X is null otherwise returns false.	The following example returns false ... | eval newField = isnull(2)

Group Aliases

Group aliases enable you to specify a group name to represent multiple columns of a specific type.

Group alias	Fields
`category`	List of all category fields
`custom float`	List of all custom float fields
`domain`	List of all domain fields
`hostname`	List of all hostname columns
`id`	List of all ID columns
`ip`	List of all IP address columns
`ip6`	List of all IPv6 address columns
`label`	List of all label columns
`mac`	List of all MAC address columns
`path`	List of all path columns
`port`	List of all port columns
`timestamp` or `time`	List of all time columns (device receipt time, agent receipt time)
`uri`	List of all URI columns
`url`	List of all URL columns
`username` or `user`	List of all user columns

Field Aliases

Field aliases enable you to specify a field name by its alias.

Note: For the fields shown in the table, you can also use presentable field names, such as Agent Address. In fact, you are encouraged to use presentable names by ArcSight Investigate's suggestions.

Field	**Aliases**
agentAddress	agt agent ip
agentHostName	ahost
agentId	aid
agentMacAddress	amac agent mac
agentReceiptTime	art
agentTimeZone	atz
agentTranslatedAddress	agent translated ip
agentType	at
agentVersion	av
applicationProtocol	app protocol
baseEventCount	cnt
bytesIn	in
bytesOut	out
categoryBehavior	behavior
categoryDeviceGroup	device group
categoryObject	object
categorySignificance	significance
categoryTechnique	technique
destinationAddress	dst destination ip destinationip dst ip dest ip target ip targetip target
destinationHostName	dhost destination name

Field	Aliases
destinationMacAddress	dmac destination mac
destinationNtDomain	dntdom
destinationPort	dpt destination port dstport dest port targetport target port
destinationProcessId	dpid
destinationProcessName	dproc
destinationTranslatedAddress	destination translated ip
destinationUserId	duid
destinationUserName	duser dst user dest user destination user dst usr
destinationUserPrivileges	dpriv
deviceAction	act
deviceAddress	dvc deviceaddr deviceip device ip
deviceCustomFloatingPoint1	cfp1
deviceCustomFloatingPoint1Label	cfp1Label
deviceCustomFloatingPoint2	cfp2
deviceCustomFloatingPoint2Label	cfp2Label
deviceCustomFloatingPoint3	cfp3
deviceCustomFloatingPoint3Label	cfp3Label
deviceCustomFloatingPoint4Label	cfp4Label
deviceCustomFloatingPoint4	cfp4
deviceCustomIPv6Address1	c6a1 device custom ipv6 1
deviceCustomIPv6Address1Label	c6a1Label

Field	Aliases
deviceCustomIPv6Address2	c6a2 device custom ipv6 2
deviceCustomIPv6Address2Label	c6a2Label
deviceCustomIPv6Address3	c6a3 device custom ipv6 3
deviceCustomIPv6Address3Label	c6a3Label
deviceCustomIPv6Address4	c6a4 device custom ipv6 4
deviceCustomIPv6Address4Label	c6a4Label
deviceCustomNumber1	cn1
deviceCustomNumber1Label	cn1Label
deviceCustomNumber2	cn2
deviceCustomNumber2Label	cn2Label
deviceCustomNumber3	cn3
deviceCustomNumber3Label	cn3Label
deviceCustomString1	Cs1
deviceCustomString1Label	cs1Label
deviceCustomString2	cs2
deviceCustomString2Label	cs2Label
deviceCustomString3	cs3
deviceCustomString3Labe	cs3Label
deviceCustomString4	cs4
deviceCustomString4Label	cs4Label
deviceCustomString5	cs5
deviceCustomString5Label	cs5Label
deviceCustomString6	cs6
deviceCustomString6Label	cs6Label
deviceEventCategory	cat
deviceHostName	dvchost
deviceMacAddress	dvcmac device mac
deviceProcessId	dvcpid

Field	Aliases
deviceReceiptTime	rt
deviceTimeZone	cat
deviceHostName	dvchost
deviceMacAddress	dvcmac device mac
deviceProcessId	dvcpid
deviceReceiptTime	rt
deviceTimeZone	dtz
deviceTranslatedAddress	device translated ip
endTime	end
eventOutcome	outcome
fileName	fname
fileSize	fsize
message	msg
requestUrl	request URL
sourceAddress	src source ip sourceip src ip
sourceHostName	shost
sourceMacAddress	smac source mac
sourceNtDomain	sntdomain
sourcePort	spt srcport src port
sourceProcessId	spid
sourceProcessName	sproc
sourceTranslatedAddress	source translated ip
sourceUserId	suid

Field	Aliases
sourceUserName	suser src user source user src usr
sourceUserPrivileges	spriv
startTime	start
transportProtocol	proto

Chapter 3: Data Quality Dashboard

Data Quality Dashboard displays more in detail information about the gap between Device Receipt Time from the raw event itself, versus the time when it actually arrives in the Vertica database. It identifies the sources that cause issues with the data. Based on the information analyzed through Data Quality Dashboard, administrators can accurately mitigate the problem. This feature also provides history of your data overtime.

Data Quality Aggregation

Data Quality is calculated and aggregated every one hour including all events that arrive to the Vertica database within the same hour. For example, the aggregated information at 10:00 AM includes all data from 10:00:00.000 to 10:59:59.999, inclusive.

> **Note:** A new table will be created in the Vertica database (during a fresh install) to store Data Quality overtime with data sources information. A Cron Job will be scheduled to trigger the aggregation process at the tenth minute of every hour. For example, if a fresh install was performed at 9:15:00 AM, the Cron Job would be scheduled to execute at 10:10:00 AM and every one hour after that.
> For upgrade, only events from the last hour will be aggregated. For example, if an upgrade was performed at 9:15:00 AM, the Cron Job would be scheduled to execute at 10:10:00 AM to aggregate all events from 9:00:00.000 to 9:59:59.999 AM, inclusive. The Cron Job will run every one hour after that.
> If switching to a different Vertica database, the user would need to wait for a few minutes before accessing the Data Quality page again.

Data Quality Dashboard Categories

Data Quality is divided into nine categories. These categories represent how big the gaps are between Device Receipt Time and Normalized Event Time.

1. **Within a minute**

   This is the category for data that arrived to Vertica with less than a one minute gap.

   Formula: Normalized Event Time - Device Receipt Time between -60000 and 60000 milliseconds

2. **Hour Behind**

   This is the category for data that is configured more than one minute to one hour late.

   Formula: Normalized Event Time - Device Receipt Time between 60001 and 3600000 milliseconds

3. **Hour Ahead**

   This is the category for data that is configured more than one minute to one hour early.

   Formula: Normalized Event Time - Device Receipt Time between -3600000 and -60001 milliseconds

4. **Day Behind**

   This is the category for data that is configured more than one hour to one day late.

   Formula: Normalized Event Time - Device Receipt Time between 3600001 and 86400000 milliseconds

5. **Day Ahead**

   This is the category for data that is configured more than one hour to one day early.

   Formula: Normalized Event Time - Device Receipt Time between -86400000 and -3600001 milliseconds

6. **Week Behind**

   This is the category for data that is configured more than one day to one week (7 days) late.

   Formula: Normalized Event Time - Device Receipt Time between 86400001 and 604800000 milliseconds

7. **Week Ahead**

   This is the category for data that is configured more than one day to one week (7 days) early.

   Formula: Normalized Event Time - Device Receipt Time between -604800000 and -86400001 milliseconds

8. **Distant Past**

   This is the most critical category for data that is configured more than one week (7 days) late.

   Formula: Normalized Event Time - Device Receipt Time > 604800001 milliseconds

9. **Far Future**

   This is the most critical category for data that is configured more than one day to one week (7 days) early.

   Formula: Normalized Event Time - Device Receipt Time < -604800001 milliseconds

Data Quality Dashboard Visualizations

Date Picker Filter

This Date Picker provides options to filter the time range for the entire Data Quality Dashboard page, including built-in Quick Ranges and Custom Range. The range of “Last 7 days” is selected by default, when you enter the page. If the Cron Job has not been run yet, it's expected to see no data displayed on the charts.

Data Timeseries

This stacked area chart represents how data is distributed between the Categories by percentage over time.

Data Sources

This visualization group consists of 3 components:

- **Category Selector:** Data sources are displayed in each of the nine Data Categories above. Far Future is selected by default.
- **Top Sources:** This donut chart represents the percentages of up-to ten top data sources with the most amount of events under the selected Data Categories. By hovering over each donut piece, you can find the IP address, the hostname, and the number of events of each source. By clicking each donut piece, more details will be displayed on the Source Timeseries side chart.
- **Source Timeseries:** This bar chart shows how many events from a data source contributed to the selected Data Categories. If available, the source with the highest number of events will be displayed by default.

Chapter 4: Viewing Event Traffic in Host Profiler

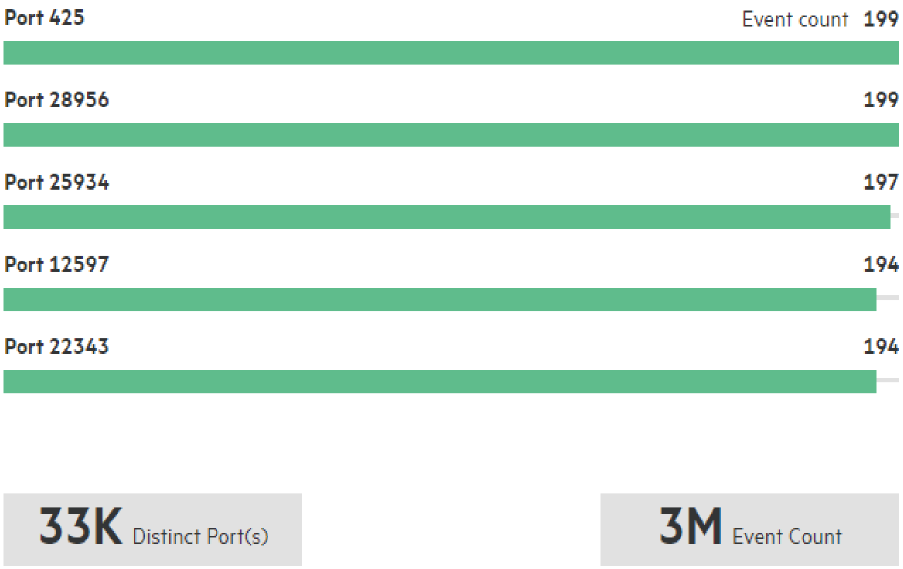
Host Profiler is a predefined dashboard where you can monitor event traffic for the five most active ports and communication paths for the host that you specify. Host Profiler provides the following visualization widgets:

- **Top 5 Outgoing Ports from the Host**

  This widget lists the five most active ports that Investigate accesses on other systems, the total number of events in the time range that you specified, and the number of total ports involved. In the following example, 983 events passed through the five most active ports, and three million events passed through thirty-three thousand ports in all.

  > **Note:** To display the table and chart click the Table Icon ▦. The table widget can display up to 100 rows. To return to the graphic, click the 📈 icon.

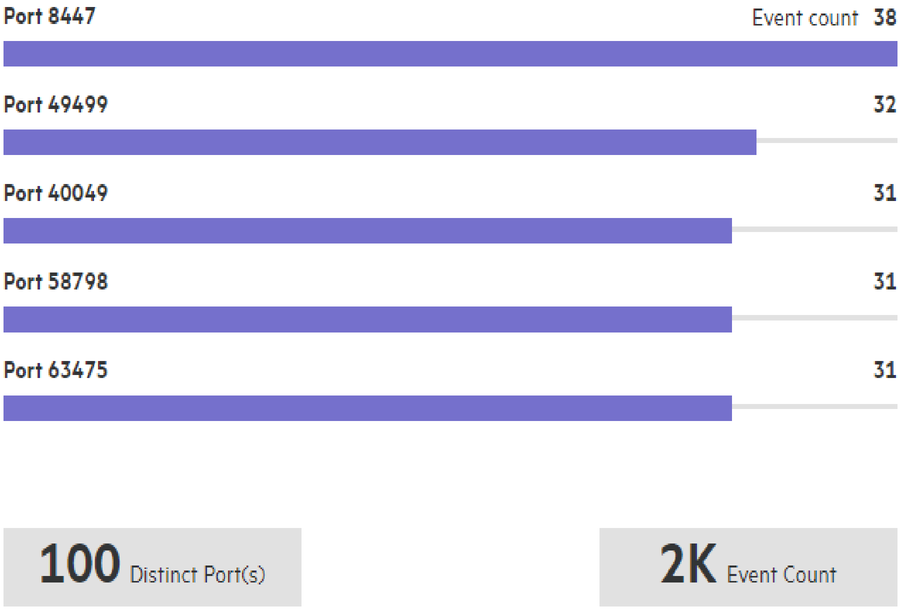
  Top 5 Outgoing Ports from the Host



- **Top 5 Incoming Ports to the Host**

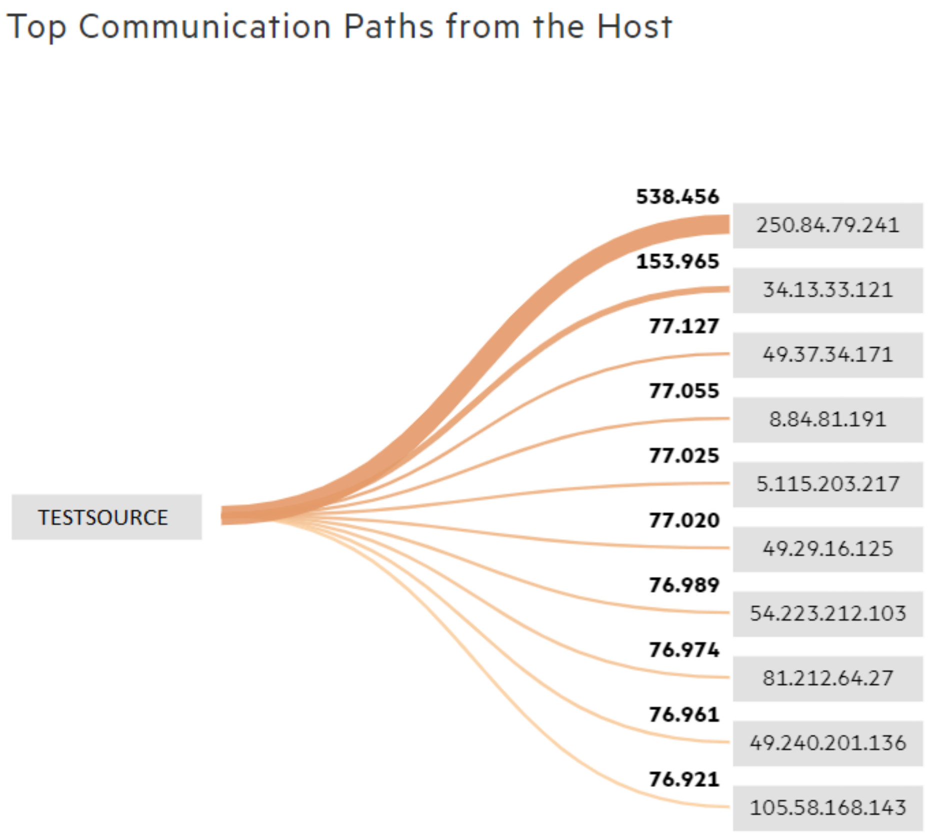
  This widget lists the five most active ports receiving events from various systems, the total number of events in the time range that you specified, and the number of total ports involved. In the following example, 163 events passed through the most active ports, and 2,000 events passed through 100 ports in all.

  Top 5 Incoming Ports to the Host



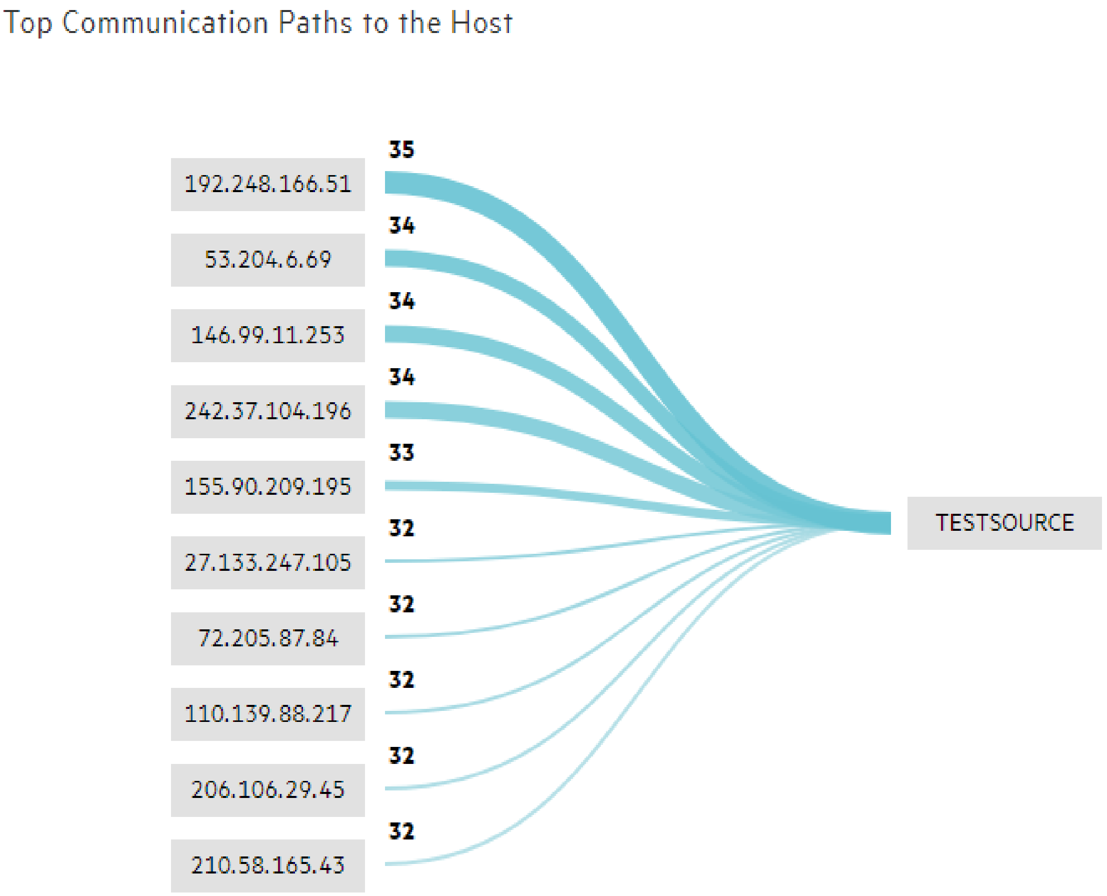
- **Top Communication Paths from the Host**

  This widget presents the ten most active systems in a Sankey chart, showing flows and their quantities in proportion to one another using the width of the lines to show their magnitudes. The number that is associated with each line is the number of events that the specified host sent.



- **Top Communication Paths to the Host**

  This widget presents the ten most active systems in a Sankey chart, showing flows and their quantities in proportion to one another using the width of the lines to show their magnitudes. The number that is associated with each line is the number of events that the specified host received.



To profile a host:

1. In the left navigation pane, select **Insights** > **Host Profiler**.
2. In the search bar, specify the IP address or name of the host that you want to profile.

   Note: Host Profiler supports IPv4 and IPv6 addresses, but not MAC addresses.

3. Select the time range for which you want to profile the host, and then click **Profile**.

   Note: Searches for events in a time range are based on the timestamps of matching events and use the time zone of the local browser. You might need to account for the time zone offset from UTC and from other time zones, including Daylight Savings Time.

   Depending on the amount of data to be profiled, the search might pause to indicate that the amount of data might impact the search performance. You might want to select a smaller time range. To resume a search, click the play button in the progress bar.

Chapter 5: Managing the Dashboard

The Dashboard allows you to add visualizations from the **Search** page in order to simultaneously monitor numerous event flows and allows you to create text box widgets for note-taking.

To create a text box widget:

1. In the left navigation pane, click **Search > Dashboard**.
2. At the top of the Dashboard page, click **Add Text Box**.

   You can begin typing in the text box.

To rename or delete a text box, click **...** in the text box panel.

To add a visualization widget:

1. In the left navigation pane, click **Search**.
2. Select the desired search.
3. Click **...** next to the desired visualization.
4. Click **Add to Dashboard**.

To view the latest data for a visualization that is associated with a real-time search, click **...**, and then click **Refresh**.

To delete a visualization from the dashboard, click **...**, and then click **Delete**. Investigate removes the widget from the Dashboard page, but does not remove the visualization from the **Search** page.

To export a dashboard widget to a PDF file, click **Export to PDF** at the top of the **Dashboard** page.

> **Note:** You cannot export individual widgets. Investigate exports all widgets on the **Dashboard** page.

For more information regarding incremental search, see "Searching Event Data" on page 5

Chapter 6: Tracking DGA Activity with DNS Analysis

A Domain Generating Algorithm (DGA) is a program or subroutine that periodically provides new domains on demand, often seen in malware. On the **Insights** > **DNS Analysis** page, Investigate provides pre-defined visualizations for displaying DGA activity.

> **Note:** To use DNS Analysis, you must use and configure the Microsoft DGA DNS Trace Log connector or Microsoft DNS Trace Log Connector on the server to send DGA events. For more information, see Configuring MS-DNS SmartConnector for DGA.

If the pre-defined visualizations do not show the desired results, you can create custom visualizations to accommodate your requirements. For more information, see Creating Custom Visualizations.

Investigate provides the following pre-defined visualizations:

- **Top Hosts by # Unique DGA Domains** shows the top hosts reporting DGA domains, sorted by the number of unique domains.
- **Top Hosts by DNS Events Sum Bytes Out** shows the top hosts reporting DNS events and the total bytes sent by each.
- **Top Unique DGA Domains by # of Hosts** shows the top DGA domains reporting DGA events, sorted by the number of reporting hosts.
- **DNS Analysis over Time** shows a graph of the number of DNS and DGA events reported over time.

Depending on the amount of data that Investigate retrieves for a visualization, the search might pause to indicate that the amount of data might impact the search performance. To resume a search, click the play button in the progress bar.

Configuring MS-DNS SmartConnector for DGA

To monitor DGA events, you must install the Microsoft DGA DNS Trace Log SmartConnector or Microsoft DNS Trace Log SmartConnector on your DNS server and configure it with the correct map files and whitelist file. The SmartConnector can be a standalone connector or ArcSight Management Center (ArcMC) can manage it. If ArcMC manages the SmartConnector, you can push the map files and whitelist files as configurations to the managed connector, as well as configure the whitelist filter. For more information, see the SmartConnector User Guide.

> **Note:** DGA configuration requires: Microsoft DGA DNS Trace Log connector 7.12.x with Parser Version: 7.12.2.8163.0 or later OR Microsoft DNS Trace Log connector >= 7.8.8070 and <= 7.9.x

To configure a standalone MS-DNS SmartConnector:

Note: This procedure shows sample content. Your files should include content that is tailored to your requirements. It is recommended to use the newer Microsoft DGA DNS Trace Log connector 7.12.x with Parser Version: 7.12.2.8162.0 or higher.

This new DGA DNS connector includes the map files by default so there is only a need to edit (as desired) the `dga_whitelist.txt` file and copy it to <ArcsightSmartConnector_Installation_ Path>/current/user/agent/ and skip to next section.

The older Microsoft DNS Trace Log connector >= 7.8.8070 and <= 7.9.x is also supported but requires the following additional steps:

For map files, adjust the sequence numbers of your new map files based on any existing map files. For example, if the last map file in the connector has the number 3, change the first new DGA map file to 4 and increment the rest of the files accordingly.

1. In a text editor, create and save the following files:

File	Description	Sample Content
dga_ whitelist.txt	White list file that includes all domains that are not scanned bythe connector DGA detection	google.com youtube.com facebook.com baidu.com wikipedia.org yahoo.com reddit.com google.co.in qq.com taobao.com amazon.com twitter.com
map.2.properties	Numbered connector mapfile tht calls the _ `domainWhitelist operation` The operation is a lookup for whitelisted domains in each event and marks them as WHITELISTED, to be dropped by the filter.	!Flags,Overwrite+ set.expr(destinationHostname).event.deviceCustomFloatingPoint2Label __domainWhitelist(destinationHostName)
map.3.properties	Numbered connector map file that calls the `dgaForbiddenTrigrams` operation The operation applies the `forbiddenTrigrams` DGA classifier on every event and returns 1 or 0 for each.	!Flags,Overwrite+ set.expr(destinationHostName).event.deviceCustomNumber1 __dgaForbiddenTrigrams(destinationHostName)
map.4.properties	Numbered connector map file that calls the `ForbiddenTrigramsHelper` operation This is a helper function that adds a label to the `dga` field in CEF.	!Flags,Overwrite+ set.expr(deviceCustomNumber1).event.deviceCustomNumber1Label __dgaForbiddenTrigramsHelper(deviceCustomNumber1)

2. Copy the whitelist file and map files to the connector you are configuring:

File	Copy to...
`dga_whitelist.txt`	`<ArcsightSmartConnector_Installation_Path>/current/user/agent/`
`x map.*.properties`	`<ArcsightSmartConnector_ Installation_ Path>/current/user/agent/map`

3. Configure the connector to filter whitelisted domains.

   The whitelist filters both exact and suffix matches.

4. Restart the Connector service.

To push map files to managed connectors:

1. In ArcMC, under **Configuration Management > All Subscriber Configurations**, create a new subscriber or subscribers (the connector or connectors that you want to configure).
2. Use the **Import** option to add the DGA map files.

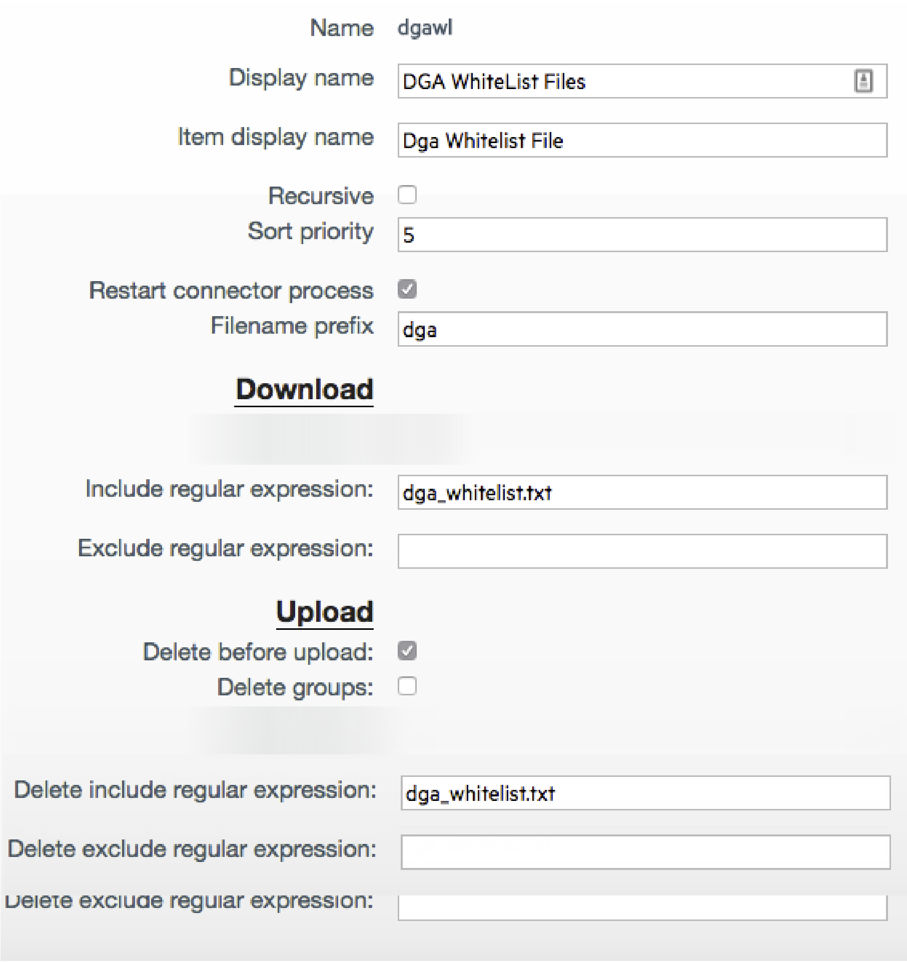
   Choose your connector and map file configuration type.
3. To add and upload your map files individually, click **Add Property**.

   Alternatively, you can create the properties files and copy their contents into the **Edit** box.

   When you are done, save your settings.
4. Use **Node Management > Containers > Restart** to restart each managed connector.

To push the whitelist file to managed connectors:

1. In ArcMC, select **Administration > Repositories**.
2. Click **Create New Repository**, and specify the settings as follows:



Name	dgawl
Display name	DGA WhiteList Files
Item display name	Dga Whitelist File
Recursive	☐
Sort priority	5
Restart connector process	☑
Filename prefix	dga
Download	
Include regular expression:	dga_whitelist.txt
Exclude regular expression:	
Upload	
Delete before upload:	☑
Delete groups:	☐
Delete include regular expression:	dga_whitelist.txt
Delete exclude regular expression:	
Delete exclude regular expression:	

3. Choose **Upload To Repository** and follow the wizard to upload the new `dga_whitelist.txt`. The wizard deletes any existing list and replaces it with the new one.

To configure the whitelist filter:

1. In ArcMC, select **Node Management > Connector** and select your MS-DNS connector.
2. Select **Runtime Parameters.**
3. Select the desired destination, such as Kafka.
4. From the parameter groups list, select `zonebasedfiltering`, and then click **Next**.
5. For Windows platforms, in the **Filter Out** field, enter `<Device_CustomFloatingPoint2_Label> = <"Device_CustomFloatingPoint2_Label> EQ WHITELISTED"`.
6. For Linux platforms, configure the whitelist filter as follows:

```
--------------------------------------------------------------------------
------
What would you like to do with
the destination?

0- Modify destination
parameters
1- Modify destination settings
2- Reregister destination
3- Add a failover destination

Please select an option:
[Modify destination settings] [0..3/back/cancel] :1

--------------------------------------------------------------------------
------
Choose a group of destination
settings to modify
.
.
```

```
.
.
10- Filters

Please select an option:
[Filters] [0..10/back/cancel] :10

---------------------------------------------------------------------------
------
Choose a group of destination
settings to modify

Filter Out[]: name EQ WHITELISTED
.
.
.

Please verify the following
parameters
Filter Out:
deviceCustomFloatingPoint2Label="deviceCustomFloatingPoint2Label EQ
Whitelisted."

Are the values correct
[yes/no/back/cancel]?YES
```

Chapter 7: Auto Pass License

This section explains the features, warnings and capacity of the Auto Pass License, as well as the steps to install the license.

Instant on License

Investigate includes an instant on license for 90 days, after this license expires, you will not be able to use the product.

Installing an Investigate term or permanent license will overwrite the instant on license.

Moving Median Events per Second (MMEPS)

MMEPS is tracked every day at GTM+0 hours, even if the license is expired or removed.

MMEPS Calculation

1. Calculate Events Per Day (EPD): Events Per Day is the total number of events ingested into Vertica database in a twenty-four hour period (for day #1 we calculate the EPD based from the time we install investigate until GTM+0 hours). The time frame is based on GTM+0 hours starting at 00:00:00 and ending at 23:59:59, regardless of any local times that may be in use.
2. Calculate Sustained EPS (SEPS): Sustained EPS is the “constant” Events Per Second that the system sustained within the twenty-four hour period(for day #1 we calculate the EPD based from the time we install investigate until GTM+0 hours). It normalizes peaks and valleys and gives a better indication of use. The formula used for this calculation is (EPD/((60*60)*24)).
3. Calculate last 45 days moving median (MMEPS): Utilizing the SEPS information recorded per day, a moving median EPS value will be identified. The Median value is calculated using last 45 day data set, and shifting the calculation window one day every twenty-four hours after the first 45 days. The official clock for calculation purposes is defined by GTM+0 hours starting at 00:00:00 to 23:59:59 regardless of local time.

Actual Calculation:

Day 1: MMEPS = SEPS of day 1

Day 2: MMEPS = AVG(SEPS of day 1 and 2)

Day 3 until last 45 days: MMEPS = median value of SEPS of day 1...45

Warnings

A warning message will be displayed in the following scenarios:

- Within thirty days before license expiration (term license or instant on license), you will receive a warning message after login indicating the license expiration date.
- Investigate will be tracking EPS every twenty four hours after installation, or when a new license is installed after the previous one expired.
- If the current calculated MMEPS exceeds license EPS capacity then there will be a warning indicating that license EPS capacity has been exceeded.
- If there are many events in TH and data ingestion to Vertica is higher than license EPS (an EPS exceed warning will be temporarily displayed until data ingestion rate normalizes).

If any of the following conditions are met you will be redirected to an invalid license page and won't be able to use the product:

- Instant on license expires.
- Investigate Term license expires.
- No license for Investigate is present.

> **Note:** In order to revert this issue, install a valid license for Investigate.

License Capacity

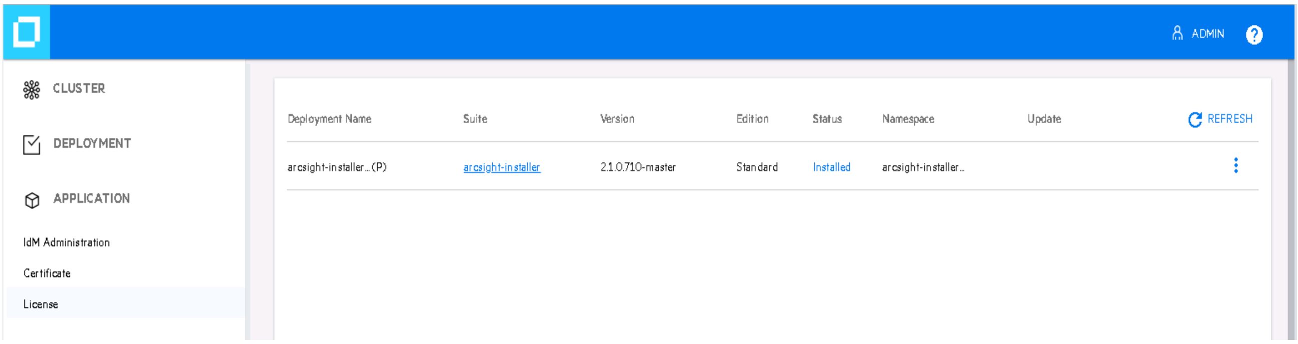
If a term or permanent license is installed, it will automatically overwrite the instant on license. License capacity will not be cumulative in this case.

If multiple licenses are installed, (term or permanent), capacity will be cumulative. Expiration date will be determined by whichever license expires first.

Installing Auto Pass License

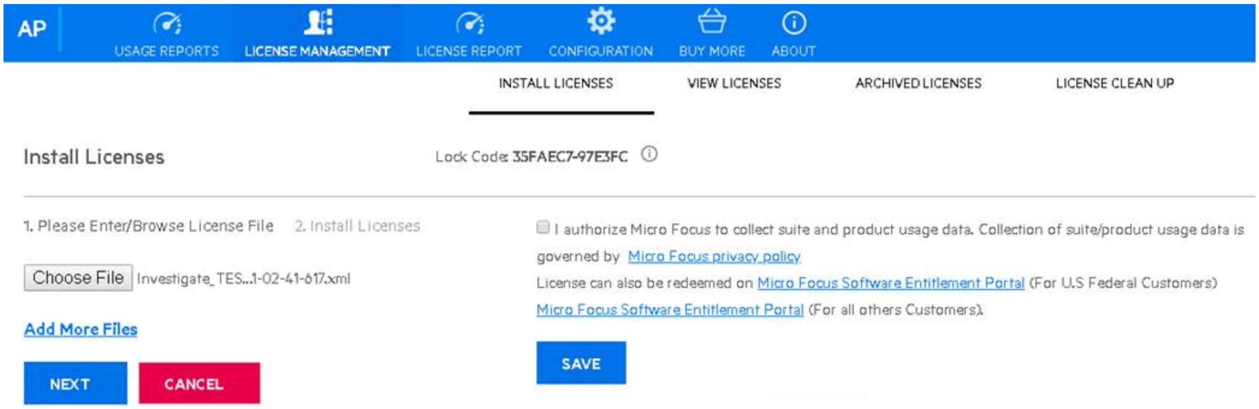
Follow the steps below to install the Auto Pass license.

1. Log in to the ArcSight Installer: `https://<Master_FQDN>:5443` or `https://<Virtualhost_FQDN>:5443` if you deployed in multi-master mode.
2. From the left menu click **Application > License.** This will display the Auto Pass license server tab.



Deployment Name	Suite	Version	Edition	Status	Namespace	Update
arcsight-installer...(P)	arcsight-installer	2.1.0.710-master	Standard	Installed	arcsight-installer...	

3. Select **Install Licenses**, click **Choose File** to upload the corresponding XML file, and click **Next.**



Install Licenses   Lock Code: 35FAEC7-97E3FC

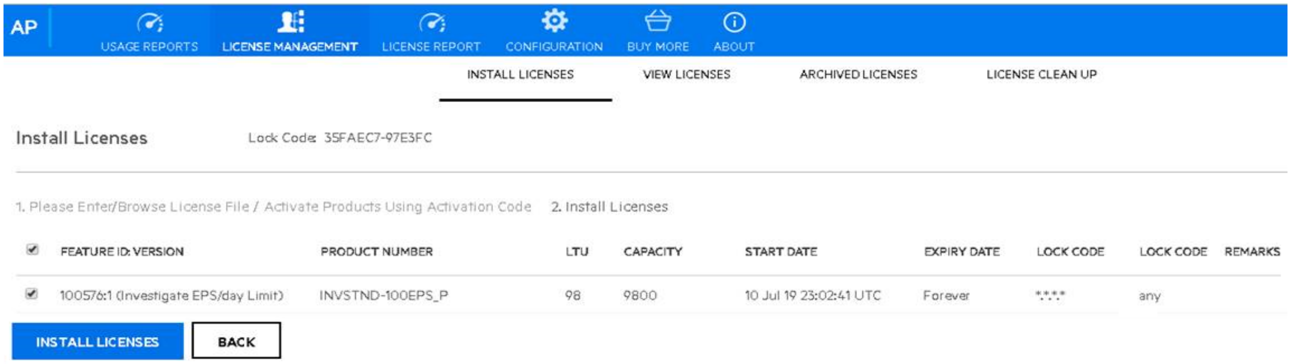
1. Please Enter/Browse License File   2. Install Licenses

Choose File Investigate_TES...1-02-41-617.xml

□ I authorize Micro Focus to collect suite and product usage data. Collection of suite/product usage data is governed by Micro Focus privacy policy

License can also be redeemed on Micro Focus Software Entitlement Portal (For U.S Federal Customers)
Micro Focus Software Entitlement Portal (For all others Customers).

4. Check the box next to the license you want to install and click **Install Licenses.**

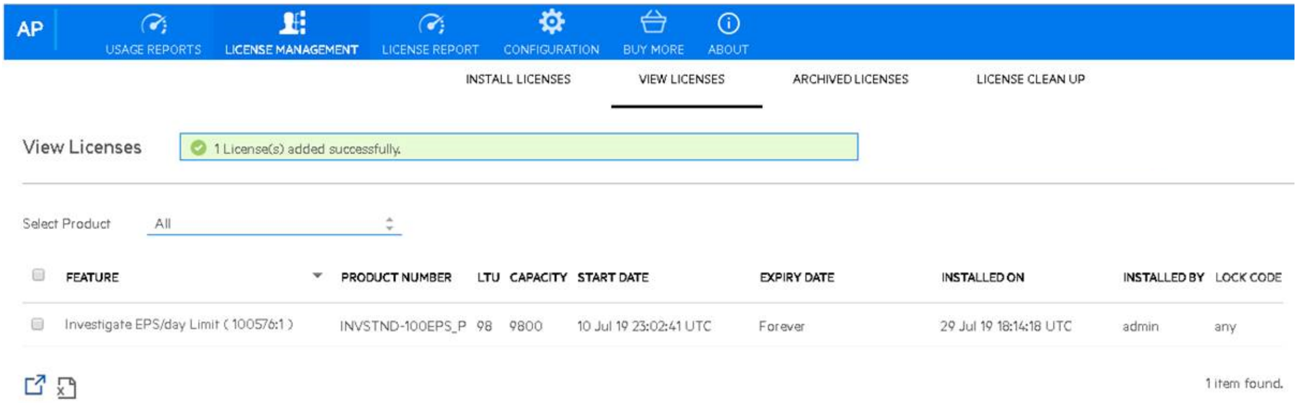


Install Licenses   Lock Code: 35FAEC7-97E3FC

1. Please Enter/Browse License File / Activate Products Using Activation Code   2. Install Licenses

☑	FEATURE ID: VERSION	PRODUCT NUMBER	LTU	CAPACITY	START DATE	EXPIRY DATE	LOCK CODE	LOCK CODE	REMARKS
☑	100576:1 (Investigate EPS/day Limit)	INVSTND-100EPS_P	98	9800	10 Jul 19 23:02:41 UTC	Forever	*.*.*.*	any	

The Auto Pass license has been added successfully.



View Licenses   ✓ 1 License(s) added successfully.

Select Product   All

□	FEATURE	PRODUCT NUMBER	LTU	CAPACITY	START DATE	EXPIRY DATE	INSTALLED ON	INSTALLED BY	LOCK CODE
□	Investigate EPS/day Limit (100576:1)	INVSTND-100EPS_P	98	9800	10 Jul 19 23:02:41 UTC	Forever	29 Jul 19 18:14:18 UTC	admin	any

1 item found.

Chapter 8: Analyzing Anomalous Data with Outlier Analytics

The Outlier Analytics feature allows you to compare incoming EventCount, BytesIn, and BytesOut values to typical values for your environment in order to identify anomalous behavior. You define and build a **model** that identifies typical behavior for your environment, and then start a scoring process that evaluates incoming events against the model. The scoring process assigns a score that indicates the degree to which the incoming data varies from the typical behavior. You view the results of the scoring process in a table that shows the top anomalous hosts, and then select a grid row to generate charts that provide additional information about the anomaly.

The model specifies a subset of data from the **Events** table that represents typical behavior on your network. When you define the model, you can specify filter criteria that identify which device behaviors you want to model. For example, you might want to look for anomalous values in events that you receive from a specific device vendor or in systems on a specific subnet.

Because the scoring algorithm is based on peer group analysis, Micro Focus recommends that you include similar devices in a model, based on activity. For example, you might want to create separate models for scoring endpoints, scoring DNS servers, and scoring databases.

For more information, see the following sections:

- Defining and Building Models
- Scoring a Model
- Viewing Scored Data

Defining and Building Models

The model defines typical EventCount, BytesIn, and BytesOut behavior for a set of IP addresses over a date range. You can specify filter criteria that identify which device behaviors you want to model.

You must have the Administrative privilege to define and build models. Each model definition applies a filter where `Source Address != NULL`.

You can define as many models as you want, but you can only build one model at a time. When you define the model, set the date range wide enough (more than 168 hours) so that the model includes a variety of device behaviors, including cyclical patterns.

When you build the model, Investigate aggregates events from the **Events** table by IP address, day of week, and hour of day for each five minute time increment, and then calculates a sum for EventCount, BytesIn, and BytesOut.

You can create and delete models, but you cannot modify them. If you want a different model, you must define and build a new one.

To define and build models:

1. From the left navigation, select **Configuration** > **Outlier**.
2. In the **Create Model Configuration** field, specify the filter criteria for building the model.

   For example, to define a specific subnet that represents a specific class of equipment (like server or data center), specify criteria similar to the following:

   `sourceAddress in subnet 10.1.1.0/24`.

   To model outbound HTTP/HTTPS traffic, specify criteria similar to the following:

   ```
   destinationPort = 80,443
   ```
3. Type over **Model Name** to specify a model name.

   > **Note:** Model names can only contain letters, numbers, and underscores. The name must start with an alpha character and cannot exceed 19 characters.

4. Specify a custom time range or select from the pre-defined ranges, and then click **Create**.

   > **Note:** Because of assumptions about the hours and days that comprise a model, do not specify a range that includes a shift in daylight savings time.

   Investigate adds the model to the **Available Models** table with a status of Created.

5. Select the model that you want to build, and then click **Build**.

   > **Note:** You can only build one model at a time.
   >
   > When you build a model, Investigate adds a lookup list of the same name to the **Configuration** > **Lookup Lists** page. You cannot view or edit this list. When you delete the model, Investigate also deletes the lookup list.

   Investigate aggregates events from the **Events** table by IP address, day of week, and hour of day for each five minute time increment, and then calculates the sum for EventCount, BytesIn, and BytesOut. It then creates conditional probability tables for sum of EventCount, sum of BytesIn, and sum of BytesOut.

6. If you want to delete a model, select the model from the **Available Models** table, and then click **Delete**.

   > **Note:** When Investigate deletes a model, it deletes the model definition and all scores that are based on that model.

Scoring a Model

After you build a model, you can start a scoring process that evaluates incoming events against the model. The process assigns a score that indicates the degree to which the incoming data varies from typical behavior. By default, Investigate selects the current date as the scoring start date. When scoring is complete, you can review the results on the **Insights** > **Outliers** page. Investigate generates a table of the top anomalous hosts, and you can select a grid row to generate charts that provide additional information about the anomaly.

You must have the Administrative privilege to score a model. You can only score one model at a time, but you can build another model while Investigate scores a model.

To score a model:

1. From the left navigation, select **Configuration** > **Outlier**.
2. From the **Available Models** table, select the model that you want to score and click **Score**.

   Note: The model must be in Build Complete status before you can score it.

3. Select the date for which you want to start the scoring process, and then click **Start**.

   Note: Because of assumptions about the hours and days that comprise a model, do not use a model that you built with daylight savings time data to score non-daylight savings time data, and do not use a model that you built with non-daylight savings time data to score daylight savings time data.

4. If you need to pause scoring because of performance or ingestion issues, click **Pause**.

   Note: If you selected a date in the past to start the scoring process, the scoring job runs frequently to catch up to the current date. To allow any running scoring jobs to complete, wait 15 minutes before performing any other action such as deleting a model or resetting scoring.

   When you are ready, click **Reset** to restart the scoring process or click **Resume** to resume the scoring process from the point at which you paused it.

When scoring is complete, use the **Insights** > **Outliers** page to view the scored data.

Viewing Scored Data

Note: The Auto-complete functionality is temporarily unavailable in search input. The following columns are available for outliers filtering in the search box:

- Source Address of <Model_Name>
- Base Event Count Score of <Model_Name>
- Bytes Out of <Model_Name>
- Bytes In of <Model_Name>

<Model_Name> corresponds to the model name being scored.

When scoring is complete, use the **Insights** > **Outliers** page to view the scored data.

After you specify search criteria for the data that you want to view, Investigate displays the top anomalous hosts that meet the criteria. When you select a host from the **Top Anomalous Hosts** table, Investigate generates charts that provide more information about the anomaly scores. You can export the charts to a PDF file. The following charts are available:

- **Outlier Scores History** compares anomaly scores of the top anomalous hosts for one week. To view details about the score for a specific date and hour, hover over the corresponding area in the chart.

  This chart is useful if you suspect a lateral attack.

- **Selected Anomalous IP** shows the anomaly score for the host that you selected over two weeks. To view details about a data point, hover over it.

  If you suspect that a host is under attack (for example, from exfiltration malware), use the **Selected Anomalous IP** chart to study the behavior of the IP over time and identify anomalous patterns.

- **Selected Anomaly Hour** compares the anomaly score for the host that you selected to the top 30 hosts for the anomaly hour. To view more details, hover over a bar in the chart, click and drag to move within the chart, and double click to reset it to its default view.

  If you suspect that a network is under attack (for example, a denial of service attack), use the **Selected Anomaly Hour** chart to study the behavior of other top 30 hosts during the anomaly hour.

After you view the outlier data, you can use the actions that are available from the grid rows in the **Top Anomalous Hosts** table to further investigate anomalies. The following actions are available:

- **Search for <IP_Address>** searches events for the host and time range for which you selected to view scoring data and displays the results on the **Search** page.
- **Profile <IP_Address>** profiles the host that you selected and displays the results on the **Host Profiler** page.
- **Filter out <IP_Adress>** displays a filter that allows you to remove the host from the **Top**

Anomalous Hosts table.

Investigate populates the filter fields based on the row from which you selected **Filter out <IP_Address>**.

To view scored data:

1. From the left navigation pane, select **Insights** > **Outliers**.
2. In the search bar, provide the following information, and then click **Detect**:
   - Select the outlier metric that you want to view (EventCount, BytesIn, or BytesOut).
   - Specify any additional query criteria that you want to apply to the data. The following parameters are currently available for filtering:
     - Base Event Count Score of <Model_Name>
     - Bytes In Score of <Model_Name>
     - Bytes Out Score of <Model_Name>
     - Source Address of <Model_Name>
     - Start Time of <Model_Name>
   - To filter a host, click a host under **Top Anomalous Hosts** and select **Filter** from the right menu.
   - Specify a custom date range for which you want to view scored data or select from the pre-defined ranges.

     > **Note:** Ensure that you specify a valid date range. Investigate displays the valid date range in the date selection area.

   > **Note:** Time range dialog displays the valid date range in the date selection area to ensure that you specify a valid date range. Scoring data is performed hourly so the time range for detection is in an hourly format (YYYY-MM-DD HH). End time hour is inclusive, e.g. If the end time is 2019-05-21 05, the scoring data from 2019-05-21 05:00-06:00 will be included. Time range dialog displays **Score Available Range** to guide users to select time range for detection.

   Investigate processes the request and generates the **Top Anomalous Hosts** table and the **Outlier Scores History**.

   > **Caution**: If Investigate retrieves a large amount of data, the search might pause. You must allow Investigate to populate the **Top Anomalous Hosts** table before you click the play button to resume the search. Otherwise, Investigate will not display any data.

3. To generate the remaining charts, select a row in the **Top Anomalous Hosts** table.
4. To perform additional actions, right-click a row in the grid and select the desired action (**Search for**

<IP_Address>, **Profile <IP_Address>**, or **Filter out <IP_Adress>**).

If you select to filter out an IP address, click **Apply** to apply the filter. To remove a filter, click **Filter** in the search bar.

5. To export the charts to a PDF file, click **Export to PDF**.

Chapter 9: Integrating SOAR Applications with Investigate

ArcSight Investigate supports integration with selected SOAR (Security Orchestration, Automation, and Response) applications. Investigate only supports integrating with one SOAR application at a time. Integrating with a second SOAR application replaces any existing integration. You can integrate Investigate with either Demisto, Operations Orchestration, or Siemplify Enterprise.

SOAR applications enable enterprises to automate their IT and security operations. When you integrate SOAR with other applications and network devices, SOAR applications can orchestrate an automatic or manual response to an IT or security event in the enterprise network. SOAR applications provide a single-pane-of-glass view for the enterprise network operations personnel.

Graphic formats called playbooks (or workflows) represent the grouping of orchestration steps and the relationships between them. When you integrate a SOAR application with Investigate, you can trigger playbooks in the SOAR application from the **Events** table in Investigate. Right-clicking an event in the **Events** table passes the entire context of the event to the SOAR application using the application’s REST API. In turn, the REST API call triggers a playbook in the SOAR application that executes the sequence of steps to accomplish the action that you initiated. When the execution is complete, the final step in the playbook sends the results of the execution to Investigate using an API exposed by Investigate, and Investigate displays the results in **Integrations** > **SOAR Notification**.

Configuring Investigate for SOAR Integration

Before you configure Investigate for SOAR integration, complete the following tasks:

1. Create the necessary playbook triggers and playbooks in the SOAR application.

   For more information, see the SOAR application documentation.
2. (Optional) Install and run the SOAR application proxy.

   In environments where a firewall exists between Investigate and the SOAR application, you must install a proxy provided by the SOAR application in the network that hosts Investigate. For more information, see the SOAR application documentation.

To configure Investigate for SOAR integration:

In Investigate, select **Configuration** > **SOAR** and set the following parameters:

Configuration Parameter	Description	Value Constraints
Application Name	The application that you want to integrate	Required
Hostname/IP	Host name or IP in URL format	Required Must start with http/https and cannot contain an end '/'
Port Number	Application port number	Two or more digits, no leading zeroes
Proxy	Proxy server URL, if required	Must start with http/https and cannot contain an end '/'
API/Application Key	Key provided by the application to access its REST APIs Use the SOAR application to generate the key. For more information, see the SOAR application documentation.	
API URI	REST end point to use to trigger a playbook in the SOAR application	Required
Login URI	REST end point to use to log in to the SOAR application	Must start with '/' and cannot contain an end '/'
Flow URI	For Operations Orchestration only, REST end point to use to get flows in Operations Orchestration	Must start with '/' and cannot contain an end '/'
Application Username	User name to log in to the SOAR application	Requires the Execute Search permission
Application Password	Password that corresponds to the user name	
Supported Actions	Comma-separated list of actions that the SOAR application supports The strings that you specify will appear in the right-click cell context menu in the **Events** table.	Must match the corresponding playbook triggers in the SOAR application
Verify Application Certificate	Enables or disables verification of server certificates from the SOAR application	
Configuration Enabled/Disabled	Enables or disables the configuration When you disable a configuration, Investigate does not display the supported actions in the right-click cell context menu in the **Events** table.	

Invoking SOAR Actions from Investigate and Viewing Results

After you configure Investigate to integrate with a SOAR application, use Investigate to invoke the actions that you configured.

First, perform a search in Investigate. In the **Events** table, right-click an event and select an action from the context menu.

> **Note:** Because Investigate does not associate actions that you configure for a SOAR application with a data type, all actions are available for all data types. For meaningful results, only right-click actions that are appropriate for the column (data type).

SOAR actions that you invoke from the **Events** table have an entry in the **SOAR Notification Details** table. To view the entries, in the left navigation pane, select **Integration** > **SOAR Notification**.

The **Status** column reflects the current status of the action:

- A Running status indicates that the playbook has not completed execution.
- A Done status indicates that the playbook execution is complete.

To view the results of the action, expand the row in the **Notification Details** table.

To delete an old entry, select the checkbox, and then click **Delete**.

Following is a snippet of Python code that shows the Investigate API and how it is used to send the results of playbook execution back to Investigate. The code is either part of an Integration script in the SOAR application that establishes the interface between that application and Investigate, or the last step in the playbook that consolidates and sends the results to Investigate.

```
import requests
import os
# First remove the proxies, since this script will
# most likely execute inside the enterprise network
if 'http_proxy' in os.environ:
del os.environ['http_proxy']
if 'https_proxy' in os.environ:
del os.environ['https_proxy']
# These arguments must be passed as arguments to the script itself
# or as arguments to a method that will be invoked to send the
# results to Arcsight Investigate
host = <must be passed as argument>
protocol = <must be passed as argument>
username = <must be passed as argument>
password = <must be passed as argument>
# Action output is the info that the SOAR application wants to
# send to Arcsight Investigate. This could be the cumulative
# output of the previous steps in the playbook
actionOutput = <must be passed as argument>
##### IMPORTANT ######
# jobid is a string that will be sent by Arcsight Investigate
# in the original request that triggered the playbook. The SOAR application
# MUST send this ID back in the results for Arcsight Investigate to
# match the result with the right action.
jobid = <must be passed as argument>
```

```
if username and password:

# first login using the arcsight Investigate credentials

# to get the SESSIONTOKEN cookie

loginRes = requests.post("%s://%s/mgmt/login" % (protocol, host),

data={"username": username, "password": password},

verify=False)

# then call the API to send the results of the playbook execution

# with the SESSIONTOKEN cookie set

result = requests.put("%s://%s/api/soarAction/update/%s" % (protocol, host,
jobid),

# jobid and actionOutput are mapped as follows

# both the keys are case-sensitive

json = {"jobid": jobid, "actionOutput": actionOutput},

cookies={"SESSIONTOKEN": loginRes.cookies.get('SESSIONTOKEN') or None},

verify=False)
```

Chapter 10: Managing Users

Investigate employs role-based access control, a method for regulating an individual user's access according to his or her role. This chapter includes information on managing users and roles in Investigate.

Managing users

Typically, users are managed in groups. When a user is created in Investigate, the user is automatically added to the All Users group (built-in group). You do not have to create additional groups in order to manage users.

The main user management tasks include:

- Creating a user:
  - Individually, as described in "Creating a User" below
  - Importing multiple users, as described in "Importing Users from ESM" on the next page
- "Adding Users to a Group" on page 70
- "Assigning a Role to a User" on page 77
- "Managing a User's Account" on page 73

Creating a User

About

Required permission:

- Create Users

  Users are identified by their email. You can only create one user with a specific email.

  By default, all users are assigned to the **All Users** group. Users cannot be removed from this group.

Procedure

Location: Left navigation > Admin > Account Groups

1. In the top navigation pane, click **Create User**.
2. In the **Create User** dialog box, enter the following user details:
   - Email
   - First name
   - Last name

3. In the **Groups** section, select the groups to which you want to add this user.

   > **Note:** We recommend that you assign the user to a group that you manage, otherwise, you won't be able to manage the user.

4. In the **Roles** section, select the roles that you want to assign to this user.

   You can only assign roles that you have yourself. All other roles are displayed as read-only.

   > **Note:** If you are creating an Admin user, in addition to the Admin role, make sure to assign the user roles that he can assign to the users that he creates. For example, the Guest and the User roles.

5. Click **Save** or **Save and Add Another** to create another user.

   The new user receives a welcome email that includes a link for creating a password for Investigate. The link is valid for 24 hours.

The groups and roles remain selected for the next user that you create.

Importing Users from ESM

About

To import users from an Existing instance of ESM in your organization, Arcsight Investigate allows the System Administrator to import users from your ESM instance.

System Admin role is required to import users from the Investigate user interface.

Users are imported into the selected Investigate groups specified during the process.

The following fields are imported:

- First name
- Last name
- Email
- Group

Users and groups are imported from the Users resource.

Import logic

- A user is identified by his or her email.
- Only users that adhere to the following conditions are imported:
  - Must have a valid email address.
  - Must be either a Web User or Normal User (user type).
  - Must be log in enabled.

- If a user does not have a first or last name in ESM, then the name is extracted from the user's email address. For example: if the user's email is ***johndoe***@microfocus.com, then the first and last name will both be ***johndoe***. You can edit the user's name in ESM before you import or after you import, in Investigate.
- All groups are imported, including deprecated groups. Deprecated groups are marked as deprecated in the import report.
- You can re-import from ESM as many times as you want. If a user has already been imported, his or her details will not be altered on re-import. However, if the user belongs to an additional group, then the user will be added to that group.
- If there are two groups with the same name, the name of one of the groups is concatenated with the parent group name.
- Empty groups are not created. If all the users in the group are invalid, the group is not created.

Once the import is complete, a report is displayed with all the users that were imported or that could not be imported and the groups that were created. You can click on the user's name to open the user's personal details. Click the Back button in the browser to return to the report.

Prerequisites

- ESM host name or IP address (as specified in the ESM Manager certificate)
- Port
- ESM credentials: Admin User ID and Password

Import is supported for ESM version 6.11 and higher.

If you have more than one ESM, perform the following procedure for each ESM server.

Procedure

Location: Left navigation > Admin > Account Groups

1. From the top navigation pane, click **Import Users**.
2. Enter the following information:
   - ESM Host Name/IP
   - Port (default 8443)
   - Admin User ID
   - Password
3. Select the roles that you want to assign to the imported users.

   You can edit roles for specific users after they are imported.
4. Click **Import Users**.

   A report is generated and displayed.

You can click on the user's name to open the user's personal details and make any required changes. Click the back button in the browser to return to the report.

New users receive a welcome email that includes a link for creating a password for Investigate. The link is valid for 24 hours.

5. If the users or groups in ESM are modified, then you can click **Start New Import** to import again.

   When importing for a second time, changes made in ESM such as deleting a group or removing a user from a group, does not affect the groups and users in Investigate.

Managing User Groups

This section includes the following topics:

What Is a User Group?

A user group is a collection of users managed by the same person. A group can have more than one manager.

Managing users in a group instead of individually, is a practical way for system administrators to delegate managerial capabilities to other administrators in the organization. For example, a system administrator can create several user groups and assign these groups to specific managers. The new managers can manage the users in their groups without having to actually create them.

In addition, managing users in groups provides managerial permissions to specific users only. This is another way of defining a clear scope of capabilities to users and limiting the span of control for admins, other than roles.

> **Note:** Groups and roles are not directly connected. Roles are assigned to a user and not to a group.

New users are automatically added to the All Users group. Users cannot be removed from this group. By default, the system administrator (or system administrators if your organization has more than one) is the manager of this group.

In order to manage a user, that user must belong to one of your groups. Even users that you create yourself must belong to one of your groups in order for you to manage them.

As a group manager, you can perform the following tasks:

- "Managing a User's Account" on page 73
- "Adding or Removing a Group Manager" on page 72
- "Terminating processes" on page 80

> **Tip:** If you do not want to create groups, you can make all the admins managers of the All Users group. This way, they can perform all possible tasks on all the users, making groups redundant. For

more information, see "Adding or Removing a Group Manager" on page 72.

Creating a User Group

Required permissions:

- Manage Groups

When you create a group you automatically become its manager.

Procedure

Location: Left navigation > Admin > Account Groups

1. In the top navigation pane, click **Create Group**.
2. In the text box in the title bar, enter a group name, and then press **ENTER**.

   The group is created.
3. To edit a group name, click on the group name to make it editable, and enter a new name.
4. "Adding Users to a Group" below.

Adding Users to a Group

Required permissions:

- Assign Users to Groups or Manage Groups

You can add users that you manage to multiple groups. They can be groups that you manage or groups that someone else manages. Once you add a user to a group that someone else manages, that person becomes the user's manager as well. You can also create new users and add them to a group.

You can add users to groups in the following ways:

- From within the group, to that specific group.

  The group that is open in the UI.
- Add users to any group.

  To any group in the system, including ones that you don't manage.
- From the user details page.

Procedure

From within the group:

Location: Left navigation > Admin > Account Groups

1. Select the group to which you want to add users.
2. Click **Add Users to this Group**.

3. In the **Add Users to this Group** dialog box, select the users that you want to add to the group, and then click **Add**.

   Only users that you manage (that belong to a group that you manage) are displayed.

4. To add a new user, click **Create New User** and follow the instructions in "Creating a User" on page 66.

From any group:

Location: Left navigation > Admin > Account Groups

1. Select the group that includes the users that you want to add to another group. The All Users group includes all the users in the system.
2. Select the users from the list, and then click **Add Users to Another Group**.
3. In the **Add Users to Another Group**, select the group or groups, and then click **OK**.

From the user details page:

1. Search for the user, as described in "Searching For a User" on the next page.
2. Click the **Groups** tab, and then click **Add/Remove User from Groups**.
3. Select the groups to which to add the user, and then click **Save**.

Removing a User From a Group

Required permissions:

- Assign Users to Groups or Manage Groups

You can remove users that you manage from any group to which they belong.

If there are users that belong only to this group (aside from the All Users group), we recommend that you move them to another group. Otherwise, these users will not be managed under any group aside from the All Users group.

You can add users to groups in the following ways:

- From the group.
- From the User Details page.

Procedure

From the group:

Location: **Left navigation > Admin > Account Groups**

1. Select the group from which you want to remove the users.
2. Select the users that you want to remove.
3. Click **Remove Users from this Group**.

From the User Details page:

1. Search for the user, as described in "Searching For a User" below.
2. Click the **Groups** tab, and then click **Add/Remove User from Groups**.
3. Select the groups from which you want to remove the user, and then click **Save**.

Adding or Removing a Group Manager

Required permissions:

- Manage Groups

If you are the group manager, you can assign additional managers to a group. Multiple managers for a group can help delegate administrative capabilities, such as managing a user's account.

Procedure

Location: **Left navigation > Admin > Account Groups**

1. Click on the group to which you want to add managers.
2. In the top navigation, click **Add/Remove Group Managers**.
3. In the **Add/Remove Group Managers** dialog box, select or remove managers, and then click **OK**.

Deleting a Group

Required permissions:

- Manage Group

You can delete a group if you are a system administrator or if you are a manager of the group.

If there are users that belong only to this group (aside from the All Users group), we recommend that you move them to another group before you delete the group. Otherwise, these users will not be managed under any group aside from the All Users group.

Procedure

Location: **Left navigation > Admin > Account Groups**

1. Select the group that you want to delete.
2. In the top navigation pane, click **Delete Group**.

Searching For a User

Required permission:

- View Users

Procedure

Location: **Left navigation > Admin > Account Groups**

1. In the top navigation, click **Search Users**.
2. In the **Search Users** dialog, enter one of the following identifiers:
   - Name
   - Email
   - ID
3. Click the user's name to open the **User Details** page.

Managing a User's Account

Required permissions:

The required permission depends on the action you want to perform. For example, if you want to reset a user's password, then you must have the Change User Password permission.

You must be the user's manager in order to manage his or her account.

What can you do?

- Edit the user's first name, last name, and email
- Reset the user's password

  When you reset a user's password the user receives a notification email automatically. The email does not include the new password. You must provide the new password to the user directly.
- Activate or deactivate a user

  A deactivated user cannot log into the system.
- Unlock a user

  Users are locked after three attempts to log into the system with the wrong credentials. When a user's account is locked, a notification is displayed on the User Details page.

Procedure

Location: **Left navigation > Admin > Account groups**

1. "Searching For a User" on the previous page.
2. To change the user's first name, last name, or email, edit the relevant fields, and then click **Save**.
3. To reset the user's password:
   a. Click **Reset Password**.
   b. On the **Reset Password** dialog box, enter a new password and confirm it.

   c. Click **Save**.

      An email is sent to the user with a notification.

4. To unlock a user account, click **Unlock** in the notification pane.

   This button only displays for locked users.

Managing Roles

Managing roles includes creating, editing, and deleting roles. For more information, see "Creating a Role" below, "Editing a Role" on the next page, and "Deleting a Role" on page 76 respectively. To manage roles, you must have the Manage Roles permission.

A user must have at least one role with one functional permission in order to log into the system.

Investigate includes built-in roles. To learn more, see "Built-in Roles" on page 78.

You can see all your roles and permissions at the following location: **My Profile > Roles & Permissions.**

What is a Role?

A role is a collection of permissions to perform certain operations or to access certain data (fields) in the system. Examples of permissions for performing operations include: View Users, Execute Search, and Manage Roles. Data access means that when you search for events, some fields are visible to you while other may not be. Restricted fields also affect visualizations; you will not be able to view charts that include a restricted field. By default, all fields are accessible. You can restrict fields when you create a role, through the Data Access tab.

Roles are created for various job functions and are assigned to users according to their function in the organization. For example, a level 1 analyst can perform operations on dashboards and searches, but cannot create new users in the system. Investigate includes a number of built-in roles. For more information, see "Built-in Roles" on page 78.

Permissions cannot be assigned directly to users.

For the full permissions' list, see "Processes" on page 80.

Creating a Role

Required permissions:

- Manage Roles

When you create a new role, it is assigned to you automatically.

To learn more about roles, see "What is a Role?" above.

Procedure

Location: **Left navigation > Admin > Roles**

1. Click **Create Role** in the top right.
2. In the text box in the title bar, enter a name for the role, and then press **ENTER**.
3. To edit a role name, click on the role name to make it editable, and enter a new name.
4. In the **Permissions** tab, select the permissions that you want to assign to this role.

   You can only assign permissions that you have yourself.
5. To restrict data access for this role, do the following:
   1. Click the **Data Access** tab.

      By default, all fields are allowed.
   2. Clear the check box from fields that you want to restrict.

      You can filter fields in the following ways:

      - Click **Restricted Fields Only**. Only restricted fields are displayed.
      - Enter a string or phrase in the **Filter** box.

Editing a Role

Required permissions:

- Manage Roles permission, as well as the role itself.

When you edit a role you can add or remove permissions, and add or remove data access. To learn more, see "What is a Role?" on the previous page

> **Note:** You cannot edit the System Admin role. For more information, see "Built-in Roles" on page 78.

Procedure

Location: **Left navigation > Admin > Roles**

1. Click the role that you want to edit.
2. To edit the role name, click the role name to make it editable and enter the new name.
3. To edit permissions, in the **Permissions** tab, select or clear the check box of the permissions that you want to add or remove from this role.

   You can only select a permission that you have yourself.
4. To edit data access, click the **Data Access** tab, and then select or clear the check box of the fields that you want to add or remove from this role.

   You can only select fields that you have yourself.

Deleting a Role

Required permissions:

- Manage Roles permission, as well as the role itself.

Note: You cannot delete the System Admin role. For more information, see "Built-in Roles" on page 78.

Procedure

Location: **Left navigation > Admin > Roles**

1. Click the role that you want to delete.
2. Click **Delete Role** in the top right.

Removing a Role From a User

About

- Manage Roles / Assign Roles to Users
  - You do not have to be the user's manager to remove his or her role; you only need the appropriate permissions.
  - If you are not a group manager, you can only assign or remove roles that you have.
  - If you are a group manager, you can remove any role that the user has but can assign only roles that you have yourself.
- You can remove roles for single and multiple users.

  The latter is recommended when to remove roles from users from the same group.

Procedure

For a single user:

Location: **Left navigation > Account Groups**

1. "Searching For a User" on page 72.
2. In the **User Details** page, click the **Roles & Permissions** tab.
3. In the **Roles & Permissions** tab, click **Assign/Remove Roles**.
4. Remove roles for this user, and then click **OK**.

Note: You can also assign or remove roles from multiple users from the User Group page.

For multiple users:

Location: **Left navigation > Account Groups**

1. On the group page, select the users from which you want to remove roles.
2. Click **Remove Roles from Users**.

   > **Note:** Only roles that you have yourself are displayed.

3. Select the roles that you want to remove, and then click **Save**.

Assigning a Role to a User

Required permissions:

- Any user with the Manage Roles or Assign Roles to Users permission can assign a role to a user in the system. You do not have to manage the user.

You can only assign roles that you have yourself.

> **Note:** If you are creating an Admin user, in addition to the Admin role, make sure to assign the user roles that he can assign to the users that he creates. For example, the Guest and the User roles.

You can assign roles in the following ways:

- From the role page.

  Recommended when you want to assign a specific role to multiple users.
- From the User Details page.

  Recommended when you want to assign multiple roles to a specific user.
- From a specific group.

  Recommended when you want to assign multiple roles to multiple users in a certain group.

Procedure

From the Roles page:

Location: **Left navigation > Admin > Roles**

1. Click the role that you want to assign.
2. In the role page, click the **Users** tab.
3. In the **Users** tab, click **Assign Role to Users**.
4. In the **Assign Role to Users** dialog box, select the users to which you want to assign the role, and then click **Save**.

From the User Details page:

Location: **User Details > Roles & Permissions**

1. Search for the user to which you want to add roles, as described in "Searching For a User" on page 72.

2. In the **User Details** page, click the **Roles & Permissions** tab.
3. Click **Assign/Remove Roles**.
4. In the **Assign/Remove Roles** dialog box, select (or clear) the roles for this user, and then click **Save**.

From a specific group:

Location: **Left navigation > Account Groups > specific group**

1. On the group page, select the users to which you want to assign roles.
2. Click **Assign Roles to Users**.
3. Select the roles that you want to add, and then click **Save**.

Built-in Roles

Investigate includes the following built-in roles:

- System Admin
- Admin
- Analyst L1
- User
- Guest

The System Admin role has the most permissions and the Guest role has the least permissions.

The System Admin role cannot be edited or deleted. If you have the System Admin role you can perform any action in the system, on any user, role, or group. A System Admin is the manager of the All Users group and cannot be removed. A system admin cannot remove his own System Admin role.

Users that have the Manage Roles permission and that have the specific role can edit or delete all the other built-in roles. By default, System Admin and Admin can edit all built-in roles.

Permissions

Note: Users must have the Execution Search permission in order to log into the system.

The following table includes the list of permissions.

Permission	Notes
View Users	A user must have this permission in order to access the Admin module. It is included in other permissions, so you do not need to actively select it if you select one of the permissions below.
Create Users	Includes: • View Users • Assign Roles to Users • Assign Users to Groups
Unlock Users	Includes View Users
Activate/Deactivate Users	Includes View Users
Change User Password	Includes View Users
Change User Email	Includes View Users
Assign Roles to Users	Includes View Users
Assign Users to Groups	Includes View Users
Manage Roles	Includes: • View Users • Assign Roles to Users
Manage Groups	Includes: • View Users • Assign Users to Groups
Manage Outlier Models and Scoring	To access the outliers feature, users need the Manage Outlier Models and Scoring permission as well as the following fields selected: • id • Bytes Out • Bytes In • Base Event Count • Source Address • Device Receipt Time
Execute Search	
Export Search Results	Includes Execute Search

Chapter 11: Processes

Processes are searches that analysts execute. Some searches are complex or are performed on a large amount of data and may take a long time to execute. Multiple searches that are in progress may slow down the performance of Investigate. In such cases, you can terminate searches that are in progress and improve performance.

> **Note:** Only processes of users that you manage display in .

Terminating processes

Procedure

Location: **Left navigation > Admin > Processes page**

- Select the desired process and then click **Stop.**

Appendix A: FAQs

Can I pin a field column in order to compare it against other field values?

In the **Events** table, allows you to pin a field column (make the column stationary) in order to better compare the column values against other columns.

For more information, see Managing Search Results Information.

Can I perform case sensitive searches?

To perform case insensitive searches, execute the below command as 'dbadmin' user in Vertica –

```
ALTER DATABASE investigate set DefaultSessionLocale = ‘en_
US@colstrength=secondary’
```

Note: Doing this may make your searches run slower.

Can I export search-results data to an Excel file?

Investigate allows you to output search results to a CSV file, which you can then import to Excel.

For more information, see Managing Search Results Information.

How much search result data can I view?

Fieldsets determine the search result fields that are visible in the **Events** table and available for creating visualizations. The default fieldset contains the most common event fields, but additional fields are available. To manage the amount of search result data that Investigate displays, you can limit the number of fields in a fieldset.

For more information, see Managing Search Results Fieldsets.

Can I view the most and least common values for a search results field?

To help filter data for security threats, you can quickly display the most and least common values for a field. For example, the **Device Vendor** field might have a top value of "bluecoat" with a count of 3,000 hits, accounting for 30 percent of 10,000 results.

For more information, see Managing Search Results Information.

Can I use SQL to specify query input?

Support for SQL statements is planned for a future release.

Can I use a SIEM with Investigate?

Currently, Investigate only supports the ArcSight Enterprise Security Manager (ESM) SIEM. Support for other SIEMs is planned for a future release.

For more information, see How ArcSight Investigate Works.

Appendix B: Debug Log Levels

You can set the log levels for each of the Investigate micro-services: search, search-engine, and user management. The following log levels are available:

- error (least verbose)
- warning
- info (default log level)
- debug
- trace (most verbose)

To change the log level:

1. Browse to `https://<installer>:5443`.
2. Navigate to suite options: **Suite > Management**
3. Click on the 3 dots at the end of the selected investigate suite and Select **Reconfigure.**
4. Under **Analytics > Log Configuration** and **Investigate > Log Configuration** select the appropriate value to update the Log Levels:

- Analytics Service Log Level
- Common Services Log Level
- Search Engine Log Level
- User Management Service Log Level
- Search Service Log Level

Send Documentation Feedback

If you have comments about this document, you can contact the documentation team by email. If an email client is configured on this computer, click the link above and an email window opens with the following information in the subject line:

Feedback on User's Guide (Investigate 3.1.0)

Just add your feedback to the email and click send.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to arcsight_doc@microfocus.com.

We appreciate your feedback!

Glossary

A

Active channel

An active channel is a tool that monitors all the activity that ESM processes for your network. An active channel displays a stream of information defined by parameters set in the active channel editor. A channel could stream events, or show the status of some resources. A channel can be further fine-tuned using in-line filters. There are three types of active channels that display different types of data: - Live Channels continuously refreshed live event data - Rules Channels display replay events for testing rules - Resource Channels display the status of certain resources, such as the assets in your network model and open cases

Aggregate data

Data that refers to numerical or non-numerical information that is (1) collected from multiple sources and/or on multiple measures, variables, or individuals and (2) compiled into data summaries or summary reports, typically for the purposes of reporting or statistical analysis—for such purposes as examining trends, making comparisons, or revealing information and insights that would not be observable when data elements are viewed in isolation.

Aggregate function

In database management an aggregate function is a function where the values of multiple rows are grouped together as input on certain criteria to form a single value of more significant meaning or measurement such as a set, a bag or a list. Common aggregate functions include : Average() and Count().

Apache Avro

Avro is a remote procedure call and data serialization framework developed within Apache's Hadoop project. It uses JSON for defining data types and protocols, and serializes data in a compact binary format. Its primary use is in Apache Hadoop, where it can provide both a serialization format for persistent data, and a wire format for communication between Hadoop nodes, and from client programs to the Hadoop services.

Apache Kafka

Apache Kafka is an open-source stream processing platform. Kafka is a distributed publish-subscribe messaging system that is designed to be fast, scalable, and durable. Like many publish-subscribe messaging systems, Kafka maintains feeds of messages in topics. Producers write data to topics and consumers read from topics. Since Kafka is a distributed system, topics are partitioned and replicated across multiple nodes.

Apache Kafka broker

Each node in an Apache Kafka cluster is called a Kafka broker.

Apache Kafka topic

The container with which messages are associated. A consumer of topics pulls messages off of a Kafka topic while producers push messages into a Kafka topic.

ArcSight Command Center (ACC)

The ArcSight Command Center is a web-based user interface that enables you to perform many of the functions found in the ArcSight Console. ArcSight Command Center provides dashboards, several kinds of searches, reports, case management, notifications, and administrative functions for managing active channels, content, users, connectors, storage, archives, search filters, saved searches, peer configuration, and system logs.

ArcSight Enterprise Security Manager (ESM)

A comprehensive software solution—a SIEM that combines traditional security event monitoring with network intelligence, context correlation, anomaly detection, historical analysis tools, and automated remediation. It consolidates and normalizes data from disparate devices across your enterprise network in a centralized view.

ArcSight Logger

Logger is a log management solution that is optimized for extremely high event throughput, efficient long-term storage, and rapid data analysis. Logger receives and stores events; supports search, retrieval, and reporting; and can optionally forward selected events. Logger compresses raw data, but can always retrieve unmodified data on demand for forensics-quality litigation data.

ArcSight SmartConnector

The interface to the objects on your network that generate correlation-relevant event data. After collecting event data for ArcSight Investigate, the connectors normalize the data in two ways: normalizing values (such as severity, priority, and time zone) into a common format, and normalizing the data structure into a common schema. SmartConnectors can then filter and aggregate events to reduce the volume of events sent to the ArcSight Manager. See SmartConnector documentation for complete details.

ArcSight Transformation Hub

Transformation Hub centralizes event processing, helps you to scale your ArcSight environment, and opens up ArcSight events to ArcSight Investigate. It enables you to take advantage of scalable, high-throughput, multi-broker clusters for publishing and subscribing to event data.

Area chart

An area chart or area graph displays graphically quantitative data. It is based on the line chart. The area between the axis and line are commonly emphasized with colors, textures and hatchings. Commonly, one compares with an area chart two or more quantities.

B

Bar chart

A bar chart or bar graph is a chart or graph that presents grouped data with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. A vertical bar chart is sometimes called a Line graph.

C

CEF

Common Event Format (CEF) is an extensible, text-based, high-performance format designed to support multiple device types in the simplest manner possible. Various message syntaxes are reduced to one-matching ArcSight

Enterprise Security Manager (ESM) normalization. Specifically, CEF defines a syntax for log records comprised of a standard header and a variable extension, formatted as key-value pairs. This format contains the most relevant event information, making it easy for event consumers to parse and use them. Other standards target a single component of the security infrastructure or are designed for specific applications. These alternatives lack the ability to support today's high-performance, real-time security requirements. For Investigate, there is CEF to Avro conversion for CEF versions 0.1 and 1.0.

CIDR notation

Classless Inter-Domain Routing (CIDR) encompasses several concepts. It is based on the variable-length subnet masking (VLSM) technique allows the specification of arbitrary-length prefixes. CIDR introduced a new method of representation for IP addresses, now commonly known as CIDR notation, in which an address or routing prefix is written with a suffix indicating the number of bits of the prefix, such as 192.168.2.0/24 for IPv4, and 2001:db8::/32 for IPv6. CIDR introduced an administrative process of allocating address blocks to organizations based on their actual and short-term projected needs.

CLI

A command-line user interface (CLI), also known as a console user interface, and character user interface (CUI), is a means of interacting with a computer program where the user (or client) issues commands to the program in the form of successive lines of text (command lines). A program which handles the interface is called a command language interpreter or shell.

Cold data

Data that is accessed less frequently by an organization. Cold data is usually stored on lower performing and less expensive storage environments in-house or in the cloud.

Column chart

A column chart is a graphic representation of data. Column charts display vertical bars going across the chart horizontally, with the values axis being displayed on the left side of the chart.

Connector

An integration element to a certain software, device format, appliance, or function through use of the product. An Onboard Connector means software that resides on the Micro Focus ArcSight appliance that communicates with other software data center. A Remote Connector is software that resides on a different computer that communicates with the Micro Focus ArcSight appliance.

Containers as a Service (CaaS)

To deliver the consistent experience for developers and IT ops, teams began using Docker for Containers as a Service (CaaS). Containers as a Service is a model where IT organizations and developers can work together to build, ship and run their applications anywhere. CaaS enables an IT secured and managed application environment consisting of content and infrastructure, from which developers are able build and deploy applications in a self service manner.

Contextual search

A form of optimizing search results based on context provided to Investigate to execute the query. For example, Investigate knows what operators to provide in the search if an IP address is specified. Likewise, if an operator is specified in the search, Investigate knows what other related operators to provide. When entering query input, Investigate can suggest fields, operators, and searches. The technology understands basic search keywords based on security terminology, database content, and user history. The search is based on such criteria as time, IPs, domains, device vendors, ports, protocols, EventCategory, and usernames. Example: Source Address = 192.10.11.12 and Destination Address less than 192.10.11.12 Investigate suggested the search items and operators.

Continuous data
Data that is not restricted to a specific value, but can occupy any value over a continuous range.

Cron job
The cron utility is a time-based job scheduler in Unix-like computer operating systems. Administrators who set up and maintain software environments use cron to schedule jobs (commands or shell scripts) to run periodically at fixed times, dates, or intervals. It typically automates system maintenance or administration—though its general-purpose nature makes it useful for things like downloading files from the Internet and downloading email at regular intervals. The origin of the name cron is from the Greek word for time, chronos.

CSV
In computing, a comma-separated values (CSV) file stores tabular data (numbers and text) in plain text. Each line of the file is a data record. Each record consists of one or more fields, separated by commas. The use of the comma as a field separator is the source of the name for this file format.

D

Data lake
A storage repository that holds a vast amount of raw data in its native format until it is needed. While a hierarchical data warehouse stores data in files or folders, a data lake uses a flat architecture to store data.

Data visualization
Data visualization is the graphical display of abstract information for two purposes: (1) Sense-making (also called data analysis) and (2) communication. Important stories live in data and data visualization is a powerful means to discover and understand these stories, and then to present them to others.

Dataset
A collection of related sets of information that is composed of separate elements but can be manipulated as a unit by a computer.

Discrete data
Data that can be numeric, such as the number of apples. It can also be categorical, like red or blue, or male or female, or good or bad.

Docker
Docker containers wrap a piece of software in a complete filesystem that contains everything needed to run: code, runtime, system tools, system libraries – anything that can be installed on a server. This guarantees that the software will always run the same, regardless of its environment. The Docker platform leverages Docker containers to enable IT operations teams and Developer teams to build, ship and run any application, anywhere. Docker containers are based on open standards, enabling containers to run on all major Linux distributions and on Microsoft Windows -- and on top of any infrastructure. Docker creates a common framework for developers and sysadmins to work together on distributed applications.

F

Fieldset

A select group of fields that determine the field information that displays in the search results for each event that matched the search query. Investigate provides a predefined, default fieldset.

Full-text search

Searches on all the tables. If you enter a string you don't know about you just search the entire columns in all the tables.

H

Hadoop

Apache Hadoop is an open source software platform for distributed storage and distributed processing of very large data sets on computer clusters built from commodity hardware. Hadoop services provide for data storage, data processing, data access, data governance, security, and operations.

Hadoop cluster

A special type of computational cluster designed specifically for storing and analyzing huge amounts of unstructured data in a distributed computing environment.

Hadoop data lake

A data management platform comprising of one or more Hadoop clusters used principally to process and store non-relational data such as log files , Internet clickstream records, sensor data, JSON objects, images and social media posts.

HDFS

The core of Apache Hadoop consists of a storage part, known as Hadoop Distributed File System (HDFS), and a processing part which is a map-reduce programming model. Hadoop splits files into large blocks and distributes them across nodes in a cluster.

Hot data

Data that needs to be accessed frequently. It is typically business-critical information that needs to be accessed quickly and is often used by a company for quick decision making. Hot data usually resides on the fastest storage -- typically flash in hybrid or tiered storage environments.

I

Integration commands

Integration commands are a set of tools in the ESM Console that make it possible to invoke scripts and utilities from several places in the ArcSight Console, and to provide snap-in views of other applications, such as ArcSight Logger and third-party applications, within the ArcSight Console. This enables you to use the ArcSight Console as a central command hub for all security-related operations. Once integrated, the commands, tools, and applications can be launched on demand from within the Console, such as from a right-click context menu within an events grid.

IoT

The Internet of things (IoT) is the inter-networking of physical devices, vehicles (also referred to as "connected devices" and "smart devices"), buildings, and other items—embedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data.

K

Key field

A field in a record that holds unique data which identifies that record from all the other records in the file or database. Account number, product code, and customer name are typical key fields. As an identifier, each key value must be unique in each record.

Kubernetes

Commonly referred to as "K8s", this is an open-source container cluster manager originally designed by Google and donated to the Cloud Native Computing Foundation. It aims to provide a "platform for automating deployment, scaling, and operations of application containers across clusters of hosts". It usually works with the Docker container tool and coordinates between a wide cluster of hosts running Docker.

L

Line chart

A line chart or line graph is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields.

Lookup list

A CSV table.

M

MAC address

A media access control address (MAC address) of a device is a unique identifier assigned to network interfaces for communications at the data link layer of a network segment. MAC addresses are used as a network address for most IEEE 802 network technologies, including Ethernet and Wi-Fi.

Micro Focus ArcSight User Behavior Analytics (UBA)

Micro Focus ArcSight User Behavior Analytics (UBA) enables security analysts to minimize the risk and impact of cyberattacks in real time. Instead of solely focusing on events and log data, Micro Focus ArcSight UBA detects unknown threats through purpose-built security analytics by creating a baseline of normal user and entity behavior and identifying anomalies associated with users and entities as they occur. By aggregating activities and multiple indicators of compromise for users, entities, and their peer groups, Micro Focus ArcSight UBA delivers insight into the highest risk users and entities—even when credentials are legitimate.

Micro Focus Security ArcSight DNS Malware Analytics (DMA)

DMA is a scalable, cloud-based threat detector that monitors DNS traffic and rapidly identifies an infected system, enabling immediate remediation in real time. The application can function in a stand-alone configuration as well as in a Security Operations Center (SOC), using Micro Focus - Security ArcSight Enterprise Security Manager (ESM) as the Security Information and Event Management (SIEM) tool.

Microservices

Microservices is a specialisation of and implementation approach for service-oriented architectures (SOA) used to build flexible, independently deployable software systems. As with SOA, services in a microservice architecture (MSA) are processes that communicate with each other over a network in order to fulfill a goal. Also, like SOA, these services use technology-agnostic protocols. The microservices approach is a first realization of SOA that followed the introduction of DevOps and is becoming more popular for building continuously deployed systems. In a microservices architecture, services should have a small granularity and the protocols should be lightweight. A central microservices property that appears in multiple definitions is that services should be independently deployable. The benefit of distributing different responsibilities of the system into different smaller services is that it enhances the cohesion and decreases the coupling. This makes it easier to change and add functions and qualities to the system at any time. It also allows the architecture of an individual service to emerge through continuous refactoring, and hence reduces the need for a big up-front design and allows for releasing software early and continuously.

N

Natural-language search

A set of pre-defined operators. Complex search: Two or more terms Separation operators: 1. And 2. Not 3. = 4. OR 5. Connecting to 6. Equals 7. List (src =1.1.1.1, 1.2.4.5) This is an OR. Example: src = 1.1.1.1 or src = 1.2.4.5

NOC

Network Operations Centers (NOCs) are implemented by business organizations, public utilities, universities, and government agencies that oversee complex networking environments that require high availability. NOC personnel are responsible for monitoring one or many networks for certain conditions that may require special attention to avoid degraded service. Organizations may operate more than one NOC, either to manage different networks or to provide geographic redundancy in the event of one site becoming unavailable. In addition to monitoring internal and external networks of related infrastructure, NOCs can monitor social networks to get a head-start on disruptive events.

O

OT

Operational Technology (OT) is hardware and software that detects or causes a change through the direct monitoring and/or control of physical devices, processes and events in the enterprise.

P

Pie chart

A type of graph in which a circle is divided into sectors that each represent a proportion of the whole. A pie chart can be used to show percentages of a whole, and represent percentages at a set point in time.

R

REST

REST (REpresentational State Transfer) is an architectural style, and an approach to communications that is often used in the development of Web services. The REST architectural style describes six constraints: - Uniform Interface - Stateless - Cacheable - Client-Server - Layered System - Code on Demand (optional)

ROS

The Read Optimized Store (ROS) is a highly optimized, read-oriented, disk storage structure, organized by projection. The ROS makes heavy use of compression and indexing. You can use the COPY statement DIRECT and INSERT parameters (with /*+direct*/ hint) to load data directly into the ROS. Note: Vertica allows optional spaces before and after the plus sign in direct hints (between the /* and the +).

RSS

RSS (Rich Site Summary; originally RDF Site Summary; often called Really Simple Syndication) uses a family of standard web feed formats to publish frequently updated information: blog entries, news headlines, audio, video.

Runbook

In a computer system or network, a runbook is a routine compilation of procedures and operations that the system administrator or operator carries out. System administrators in IT departments and NOCs use runbooks as a reference. Runbooks can be in either electronic or in physical book form.

S

Scatter plot chart

A scatter plot (also called a scatter graph, scatter chart, scattergram, or scatter diagram) is a type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data. If the points are color-coded, one additional variable can be displayed.

Security analyst

The primary user of ArcSight Investigate. This user relies on the overall ArcSight log collection and search capabilities for successfully triaging security incidents. Ultimately, security analysts want to get actionable insights from a search.

Security architect

This user is responsible for determining the overall ArcSight deployment and how this product fits into the SIEM architecture of the organization. This includes integration with other systems such as Hadoop which may be used for storing additional log data.

Security engineer

This user is responsible for data sources and determining how security analysts can effectively triage security incidents and security threat.

Security posture

Your overall security plan – the approach your organization takes to security, from planning to implementation. It is comprised of technical and non-technical policies, procedures and controls, that protect you from both internal and external threats.

SIEM

In the field of computer security, Security Information and Event Management (SIEM) software products and services combine Security Information Management (SIM) and Security Event Management (SEM). They provide real-time analysis of security alerts generated by network hardware and applications. Micro Focus Security ArcSight Enterprise Security Manager (ESM) is an example of a SIEM product.

SMTP

SMTP (Simple Mail Transfer Protocol) is a TCP/IP protocol used in sending and receiving e-mail.

SOC

An information security operations center ("ISOC" or "SOC") is a facility where enterprise information systems (websites, applications, databases, data centers and servers, networks, desktops and other endpoints) are monitored, assessed, and defended.

Sparkline

A small graphic designed to give a quick representation of numerical or statistical information within a piece of text, taking the form of a graph without axes.

Subnet

A logical subdivision of an IP network. Computers that belong to a subnet are addressed with a common, identical, most-significant bit-group in their IP address. This results in the logical division of an IP address into two fields, a network or routing prefix and the "rest" field or host identifier. The rest field is an identifier for a specific host or network interface.

System admin

This user is responsible for the deployment, administration and day to day operations of the product. They will need the necessary monitoring and administrative controls to ensure that the product is available and functioning with optimal performance for security analysts.

T

Text box widget

From the Dashboard, you can use this widget to create, edit, and delete a text note.

TLS

Transport Layer Security (TLS) and its predecessor, Secure Sockets Layer (SSL), both frequently referred to as "SSL", are cryptographic protocols that provide communications security over a computer network. In the case of ArcSight Investigate, TLS is used between the user management, search interface, search engine, and Vertica database modules.

Truncate table

This removes all rows from a table, but the table structure and its columns, constraints, indexes, and alike remain.

Tuple

A tuple is a sequence of immutable Python objects. Tuples are sequences, just like lists. The differences between tuples and lists are, the tuples cannot be changed unlike lists and tuples use parentheses, whereas lists use square brackets. Creating a tuple is as simple as putting different comma-separated values.

V

Vertica

An advanced SQL database that can address the most demanding Big Data analytics initiatives. It introduces a unified architecture and advanced in-database analytics capabilities that enable users to conduct sophisticated analysis at industry-leading scale and speed, regardless of where their data resides.

Z

ZooKeeper

Apache ZooKeeper is a distributed hierarchical key-value store, which is used to provide a distributed configuration service, synchronization service, and naming registry for large distributed systems. ZooKeeper was originally a sub-project of Hadoop. ZooKeeper's architecture supports high availability through redundant services. The clients can thus ask another ZooKeeper leader if the first fails to answer. ZooKeeper nodes store their data in a hierarchical name space, much like a file system or a tree data structure. Clients can read from and write to the nodes and in this way have a shared configuration service. Updates are ordered.