Micro Focus Security ArcSight Investigate

Software Version: 2.10

Deployment Guide

Document Release Date: March 30, 2018
Software Release Date: March 2018

Legal Notices

Warranty

The only warranties for products and services of Micro Focus and its affiliates and licensors (“Micro Focus”) 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. The information contained herein is subject to change without notice.

Restricted Rights Legend

Confidential computer software. Except as specifically indicated otherwise, a valid license from Micro Focus is required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright Notice

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.

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

Revision History

Date	Description
March 30, 2018	Initial release of this document.

Contents

Document Revision History

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

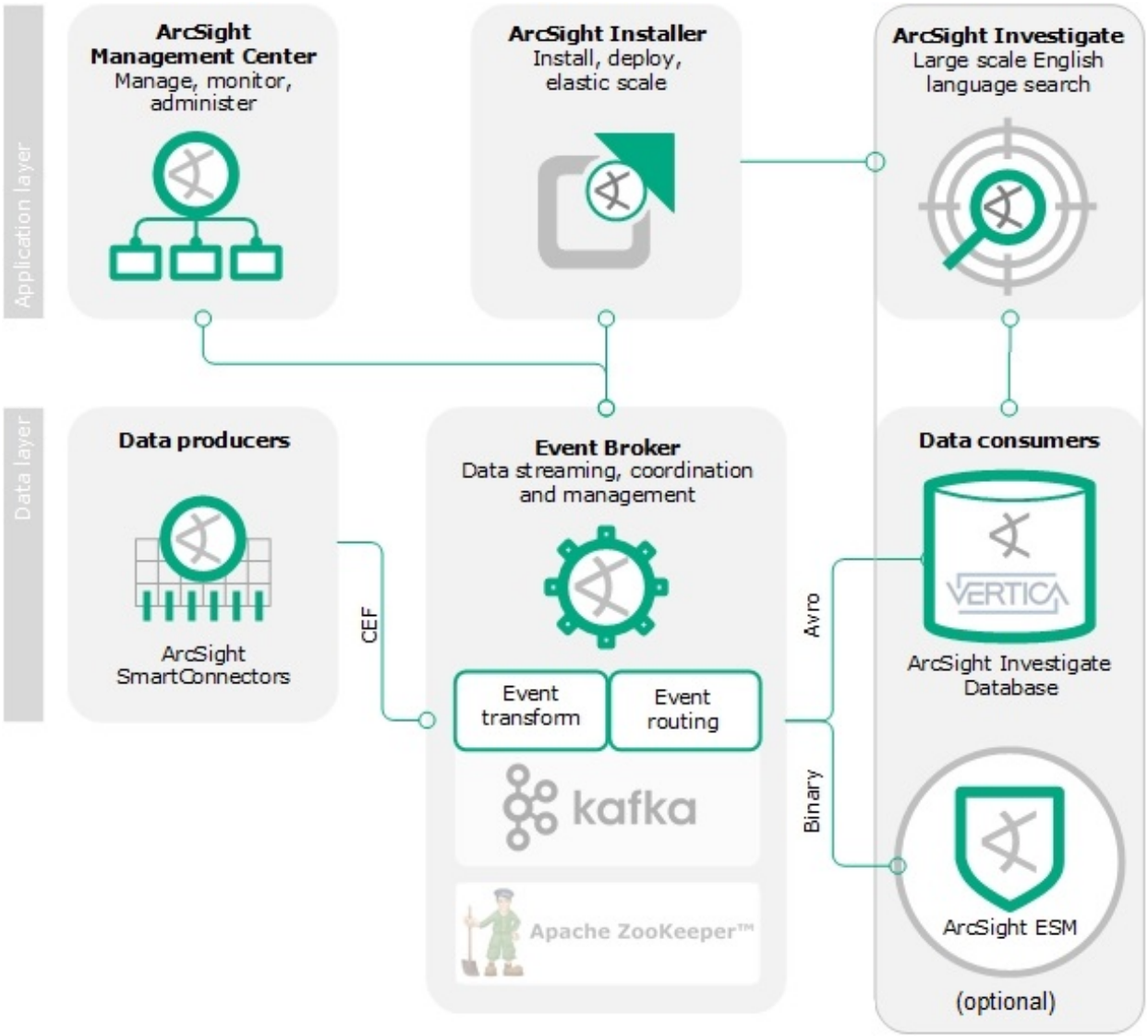
- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.

To check for recent updates or to verify that you are using the most recent edition of a document, go to the : ArcSight Product Documentation Community on Protect 724.
(missing or bad snippet)

Chapter 1: Introduction

About ArcSight Investigate

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 you can view and search on them. You can use the English-like search language to generate results from which to create reports and visualizations.

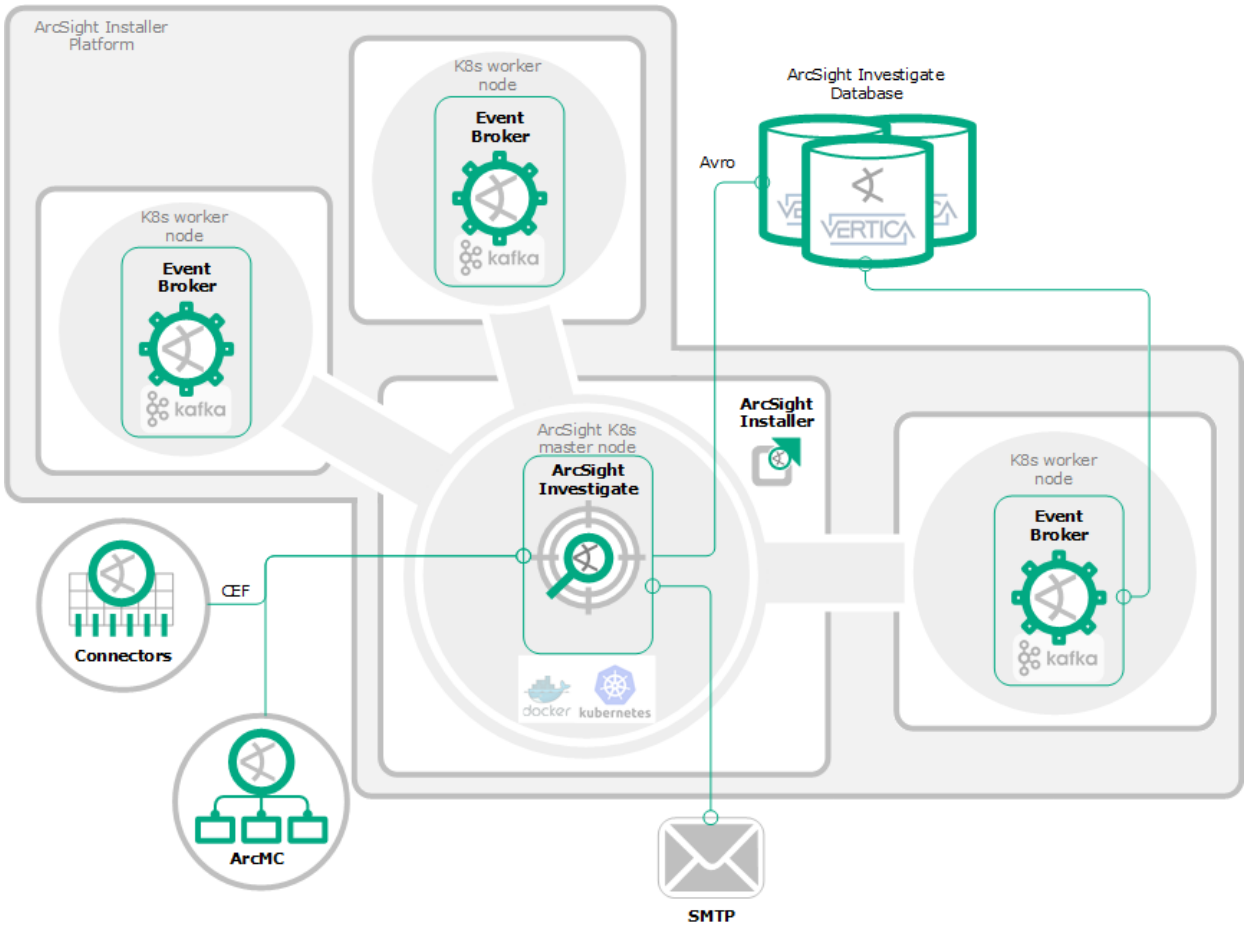


Component	Description
ArcSight Investigate	High-capacity data management, search, and analysis web application.
ArcSight Investigate Vertica database	The ArcSight Investigate analytic database powered. Vertica is installed separately.
ArcSight Installer	A web application for deploying and configuring the ArcSight Investigate components, including Investigate and Event Broker. The components are managed in a Kubernetes cluster. The master node hosts the ArcSight Installer web application and the Investigate web application, and the worker nodes host the Event Broker.
ArcSight SmartConnectors	SmartConnectors collect and normalize event data from nodes on your network. Connectors normalize event 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 system. ArcSightSmartConnectors, installed and maintained separately, are producers that publish data to Event Broker. You can subscribe to data managed by Event Broker with Investigate, ArcSight Deployment Platform (ADP) Logger, ArcSight ESM, Apache HDFS, or your own third-party consumer.
Event Broker	ArcSightEvent Broker, a product of the ADP suite, centralizes event processing, enabling you to take advantage of scalable, high-throughput, multi-broker clusters for publishing and subscribing to event data. Event Broker coordinates and manages data streams, which enables your ArcSight environment to scale, and opens up ArcSight events to third-party data solutions.
ArcMC	HPE ArcSight Management Center (ArcMC) is a centralized management tool that simplifies security policy configuration, deployment maintenance, and monitoring efficiently and cost-effectively. ArcMC provides run-time management of Event Broker topics. ArcMC is sold as part of ADP.

ArcSight Investigate deployment architecture

ArcSight Investigate runs in Docker containers managed by Kubernetes and deployed from the ArcSight Installer application. ArcSight Investigate is deployed on a node within the Kubernetes cluster. It can be installed on a master or worker node within the cluster. ArcSight recommends that for production deployments, you install three Kubernetes master nodes and three Kubernetes worker nodes. Also supported is a configuration with a single Kubernetes master node, and three Kubernetes worker nodes. In this case, the ArcSight Event Broker should be installed on the three worker nodes and ArcSight Investigate should be installed on the master node. Consult with your ArcSight Technical Specialist for architecture options.

The image below is a typical representation of the deployment architecture, with one master node and three worker nodes.



Deployment component	Host	Functional contents
ArcSight Installer Platform	Install the platform on the master node and each work node.	ArcSight Installer application
Kubernetes master node	3 VMs or physical servers	• 3 Kubernetes master nodes Multi-master deployment requires a minimum 3 nodes, while a single master node only requires 1. • Investigate • ArcSight Installer application
Kubernetes worker nodes	3 VMs or physical servers	3 Kubernetes Event Broker nodes A worker node is needed for each Event Broker instance. For multi-master deployment, 1 worker node is needed for Investigate.

Deployment component	Host	Functional contents
Vertica database	3 physical servers	One ArcSight Investigate Vertica database cluster with 3 nodes
ArcSightSmartConnectors	Stand-alone or part of ArcMC	Normalizes event data from network devices and formats as CEF.
ArcSight Management Console	Separate installation	Provides run-time management of Event Broker topics.
SMTP server	Separate installation	Provides the ability for ArcSight Investigate to send notification messages to users.

Deployment overview

Before you can deploy ArcSight Investigate deployment you must first install the ArcSight Installer, Event Broker, and the Vertica database.

Note: ArcSight recommends installing and running these components in a test environment before putting them into production.

These components require configuration after you install Investigate containers.

1. Complete the installation requirements.
   a. Ensure that you upgraded to ArcSight Installer 1.40 and Event Broker 2.20.

      See the *ArcSight Event Broker 2.20 Deployment Guide* for details.
   b. Ensure that Event Broker and Investigate each have dedicated servers.

      If other applications are running on the same server as Event Broker and Investigate, there will be a significant performance penalty and potential problems.
   c. Generate an SSH certificate on the master node in order to allow connections to the worker nodes.
      See "Generating a key pair on the master node for worker nodes" on page 1.
2. Obtain the Investigate image.
   - For online retrieval, see "Obtaining Investigate images online for the Installer" on page 28.
   - For offline retrieval, see "Obtaining Investigate images offline for the Installer" on page 28.
3. Ensure that Event Broker is installed and deployed.

   See the *ArcSight Event Broker Deployment Guide*.
4. Install and deploy all Investigate components.

   See "Deploying Investigate images" on page 29.

5. Ensure that the images have completed deployment.

   ```
   kubectl get pods --all-namespaces
   ```

   See "Deploying Investigate images" on page 29.
6. Install the Vertica database.

   See "Installing Vertica" on page 23.
7. Configure Event Broker if necessary.

   See the *ArcSight Event Broker Deployment Guide*.
8. Configure Investigate, including the Investigate Vertica database and SMTP server.).

   See "Configure ArcSight Investigate and components" on page 40

Planning your deployment

Before deploying, ensure that you have the latest version of this document, available for download from the ArcSight Product Documentation Community on Protect 724.

TLS planning

The various components in the ArcSight Investigate system interact using encrypted communication implemented using Transport Layer Security (TLS) 1.2 protocol.

TLS implementation requires digital certificates. Before you begin the installation process, you must decide on the type of certificate you prefer to use:

- A self-signed certificate. Kubernetes includes the capability to generate self-signed certificates. By default, the Kubernetes installation process generates certificates for the Kubernetes cluster, but you can instruct otherwise during the installation process. You can also generate a Kubernetes certificate for other components in the system, which require a certificate, like the ArcSight Investigate Vertica database. For more information on generating a Kubernetes certificate, see Generate signed certificates for consumers.
- A certificate signed by a certificate authority (CA). Depending on your organization's security policy, you might be required to use a certificate from a trusted CA. In this case, make sure that you have a root certificate file and a private key file. Copy these files to the designated Kubernetes master node.

Note: The certificates cannot be reconfigured after installation.

Network planning

- Ensure that each node is configured with a fully qualified domain name.
- Ensure proper DNS configuration across all systems including correct forward and reverse DNS lookups.
- Enable internet access in order to download the product container images.
- If your organization's network has a proxy, define the proxy environment variable on all servers. Define these variables per user, and not system-wide.

Plan encryption modes

Before installing Investigate and Event Broker, determine the encryption mode you want to use to encrypt communications between ArcSight components. Set up the other ArcSight components with the encryption mode you intend to use before connecting them to the Event Broker. The security mode of systems connected to Event Broker (Consumers, Producers, ArcMC) must be the same as the security mode set for Event Broker. Changing encryption modes after Event Broker has been deployed will require system down time. If you do need to change the security mode after Event Broker deployment, see the *Event Broker Administrator's Guide*.

Product	Preparations needed	Open ports	Supported encryption modes	Guidance documentation
ArcMC	Install ArcMC before ArcSight Investigate and Event Broker installation.	38080	• TLS • FIPS • ClientAuth	*ArcMC Administrator's Guide*
ArcSight SmartConnectors	ArcSight SmartConnectors and ArcMC onboard connectors can be installed and running prior to installing ArcSight Investigate and Event Broker. FIPS mode setup is not supported between Connector version 7.5 and Event Broker. TLS and ClientAuth are the only encryption methods supported between SmartConnector version 7.5 and Event Broker.	9093	• TLS • FIPS • ClientAuth	*SmartConnector User Guide* *ArcMC Administrator's Guide*

Product	Preparations needed	Open ports	Supported encryption modes	Guidance documentation
ArcSight ESM (optional)	ArcSight ESM can be installed and running prior to installing ArcSight Investigate and Event Broker. ESM ingests events faster than Investigate does. (Investigate Scheduler ingests events at 22K per second while ESM ingests events at 30K per second.) You can leave the ingestion rate asynchronous, or you can equalize them by setting the ESM ingestion rate to a lower rate at the connector so that Investigate and ESM ingest rates are closer. This will reduce the likelihood of a lag in search results on Investigate launched from ESM.	9093	• TLS • FIPS • ClientAuth	*ESM Installation Guide* *ESM Administrator's Guide*
ArcSight Logger (optional)	ArcSight Logger can be installed and running prior to installing Event Broker.	9093	• TLS • FIPS • ClientAuth	*Logger Administrator's Guide*

Chapter 2: ArcSight Investigate support matrix

Supported operating systems

Version	Component	Operating system
2.10	ArcSight Investigate	RHEL 7.3 64-bit* RHEL 7.4 64-bit CentOS 7.3 64-bit* CentOS 7.4 64-bit * Linux kernel version 3.10.0-514.26.2.el7.x86_64 (or above)
	ArcSight Investigate Vertica 8.1.1-3 database	RHEL 7.3 and CentOS 7.3

Supported browsers

Browser	Version
Microsoft Edge	Version available at the time of release.
Google Chrome	Version available at the time of release.
Mozilla Firefox	Version available at the time of release.

Supported product compatibility

Product	Version
ArcSight Event Broker	2.20
ArcSight SmartConnector	7.5 and later
ArcMC	2.80
ArcSight Logger	6.6 and later
ArcSight ESM	6.11

Chapter 3: Prerequisites for installation

ArcSight Installer and Event Broker requirements

- Ensure ArcSight Installer is at version 1.40.

  If you have an older version, see the *ArcSight Event Broker 2.20 Release Notes* for upgrade details.
- Ensure ArcSightEvent Broker is at version 2.20.

  If you have an older version, see the *ArcSight Event Broker 2.20 Release Notes* for upgrade details.

System requirements

General sizing guidelines

Provision the servers (or VMs) that you are using for the deployment, based on the general sizing guidelines provided here. This information is based on a default setup.

For tailored sizing recommendations for a production environment, contact ArcSight Customer Support.

For supported platforms and operating systems, see the ArcSightInvestigate Support Matrix.

Component	Nodes	Resources needed	Needed ports
ArcSight Investigate + Event Broker	1 master 3 worker	• One CPU with 24 cores • 32 GB RAM • 8 TB disk space • Linux kernel version 3.10.0-514.26.2 (or above) • Java (OpenJDK) 1.8.0_121 or higher • Method for obtaining Docker containers, either via Internet (or proxy) or other internal method • 10 GigE network **Note:** If you choose to deploy ArcSight Investigate on a worker node, the nginx reverse proxy used to connect to Investigate is always deployed on the master node. Therefore, no matter where Investigate is deployed in a Kubernetes cluster, you should always access Investigate using the host/IP of the master node.	Kubernetes: 2379, 2380, 4001, 4194, 5000, 5443, 8080, 8088, 8200, 8285, 8443, 10248-10252, 10255, 30001 Network File System (NFS): 111, 2049, 20048, 37189 For required Event Broker ports, see "System requirements" in the *ArcSight Event Broker Deployment Guide*. Investigate: 5443, 21085, 30001
Investigate Vertica Database and Scheduler	3	**Important:** The Vertica database must be installed on the same sub-network as the Investigate master and worker nodes. • 2 CPUs with 24 cores • 128 GB RAM • 8 TB disk space • 10 GigE network minimum (dual recommended) Recommendation: Install Vertica on a dedicated physical server. Example: HPE Proliant G9 or similar Virtual environment: HPE Vertica performs better on a physical server than in a virtualized environment because of the overhead and resource constraints imposed by the virtualization software. See HPE Vertica Analytics Platform Version 8.1.x Documentation for more information.	5433
ArcMC (part of ADP)	1	• One CPU quad-core • 16 GB RAM • 50 GB of free disk space For ArcMC deployment details, see the *ArcMC Administrator's Guide*.	
SmartConnectors (part of ADP)	1	SmartConnector version 7.5 (can be stand-alone or managed by ArcMC) For ArcSightSmartConnector deployment details, see the *SmartConnector User's Guide*.	

Network requirements

Caution: If the default network ranges specified here are in use in your network environment, the installation may fail, or random failures may be experienced after installation.

By default, the Installer uses the following network ranges:

- 172.16.0.0/16 — sub-network of 65,536 addresses for the operation of Kubernetes pods with containers running in them. Each pod operates with the /24 sub-network from following range.
- 172.30.78.0/24 — sub-network of 256 addresses for the operation of Kubernetes services, including internal Kubernetes DNS service located on pod 172.30.78.78.

For the /16 and /24 address ranges, ensure that your network is conflict free. If these address ranges are occupied and/or not accessible due to network configuration, utilize another address range by making corresponding changes to the `POD_CIDR`, `SERVICE_CIDR` and `DNS_SVC_IP` parameters in the `./<path to the secure location on master node>/arcsight-installer-<version>/arcsight-installer-master.sh` script.

Firewall requirements

The following ports need to be free and available for firewall configuration.

- **Kubernetes:** 2379,2380,3000,4001,4194,5000,5443,8080,8088,8200,8285,8443,10248-10252,10255
- **NFS:** 111,2049,20048,37189

- **Investigate:** 5443,21085,30001

ArcSight Installer configures firewall settings during setup (in case `firewalld.service` is up and running) on both Kubernetes master and Kubernetes nodes.

Vertica requires several ports to be open on the local network. Vertica does not recommend placing a firewall between nodes. (All nodes should be behind a firewall.) If you must use a firewall between nodes, ensure the following ports are available:

Port	Protocol	Service	Notes
7	TCP	Management Console	Required by Management Console to discover Vertica nodes.
22	TCP	sshd	Required by Administration Tools and the Management Console Cluster Installation wizard.
5433	TCP	Vertica	Vertica client (such as vsql, ODBC, JDBC) port.

Port	Protocol	Service	Notes
5434	TCP	Vertica	Intra- and inter-cluster communication. Vertica opens the Vertica client port +1 (5434 by default) for intra-cluster communication, such as during a plan. If the port +1 from the default client port is not available, then Vertica opens a random port for intra-cluster communication.
5433	UDP	Vertica	Vertica spread monitoring.
5444	TCP	Vertica Management Console	MC-to-node and node-to-node (agent) communications port. See Changing MC or Agent Ports.
5450	TCP	Vertica Management Console	Port used to connect to MC from a web browser and allows communication from nodes to the MC application/web server. See Connecting to Management Console.
4803	TCP	Spread	Client connections.
4803	UDP	Spread	Daemon to Daemon connections.
4804	UDP	Spread	Daemon to Daemon connections.
6543	UDP	Spread	Monitor to Daemon connection.

Configuring proxy settings

About

Comment out proxy data in the `/etc/profile.d/proxy.sh` file on all nodes, if it is being used.

If you are using a proxy server in your environment, then add your proxy data to the `~/.bashrc` file.

Procedure

Update the `.bashrc` file according to the following example:

```
export http_proxy=http://<proxyserver>:8080/

export https_proxy=http://<proxyserver>:8080/

export HTTP_PROXY=http://<proxyserver>:8080/

export HTTPS_PROXY=http://<proxyserver>:8080/

export no_proxy="<master ip>,<worker-1 ip>,<worker-2 ip>,<worker-3
ip>,localhost,<domain>"

export NO_PROXY="<master ip>,<worker-1 ip>,<worker-2 ip>,<worker-3
ip>,localhost,<domain>"
```

Chapter 4: Install the Investigate Vertica database

Prerequisites

- Ensure that Vertica has a dedicated server or servers. If other applications are running on the same server, there may be a performance penalty and potential problems.
- Ensure that your file system type is ext4.

Configuring the Vertica server

About

The server configuration described here is based on an HPE ProLiant DL380 Gen9 Server with 128 GB memory. Since your server may be different and your environment unique to you, this procedure serves as a reference.

Note: For some commands where further explanation may be necessary, there is a reference number or statement parenthetically attached. Refer to this number or statement in the "Reference" section for any necessary details.

Example: `#### Resolve WARN (S0112)`

Procedure

1. Provision Vertica server.
   - No Logical Volume Manager (LVM) partition
   - Partition type: ext4
   - Minimum 2 GB swap space
   - RHEL 7.3 (or greater) or CentOS 7.3 (or greater)
2. Add the following parameters to `/etc/sysctl.conf`. The changes take effect after rebooting.

```
## Increase number of incoming connections

net.core.somaxconn = 1024

## Sets the send socket buffer maximum size in bytes.

net.core.wmem_max = 16777216
```

```
## Sets the receive socket buffer maximum size in bytes.

net.core.rmem_max = 16777216

## Sets the receive socket buffer default size in bytes.

net.core.wmem_default = 262144

## Sets the receive socket buffer maximum size in bytes.

net.core.rmem_default = 262144

## increase the length of the processor input queue

net.core.netdev_max_backlog = 100000

net.ipv4.tcp_mem = 16777216 16777216 16777216

net.ipv4.tcp_wmem = 8192 262144 8388608

net.ipv4.tcp_rmem = 8192 262144 8388608

net.ipv4.udp_mem = 16777216 16777216 16777216

net.ipv4.udp_rmem_min = 16384

net.ipv4.udp_wmem_min = 16384

#### Increase the number of outstanding syn requests allowed.

net.ipv4.tcp_max_syn_backlog = 4096

#### Based on 128 GB of memory, (Tuning Linux Dirty Data Parameters for
Vertica)

dirty_ratio = 8

#### Resolve WARN (S0112)

vm.swappiness = 1
```

3. Add the following parameters to `/etc/rc.local`.

   The changes take effect after rebooting.

```
echo 'echo deadline > /sys/block/sda/queue/scheduler' >> /etc/rc.local
#### Resolve FAIL (S0150)

echo '/sbin/blockdev --setra 2048 /dev/sda' >> /etc/rc.local #### Resolve
FAIL (S0020)

echo '/sbin/blockdev --setra 2048 /dev/sdb' >> /etc/rc.local #### Resolve
FAIL (S0020)
```

```
echo 'cpupower frequency-set --governor performance' >> /etc/rc.local ####
CentOS only, resolve WARN (S0141)

chmod +x /etc/rc.local
```

4. Increase the process limit.

   ```
   /etc/security/limits.d/20-nproc.conf
   ```

   add:

   ```
   * soft nproc 10240

   * hard nproc 10240

   * soft nofile 65536

   * hard nofile 65536

   * soft core unlimited

   * hard core unlimited
   ```

5. Set `max_cstate`.

   ```
   vi /etc/default/grub
   ```

   Append line `GRUB_CMDLINE_LINUX` with `intel_idle.max_cstate=0 processor.max_cstate=1`
   Example:

   ```
   GRUB_CMDLINE_LINUX="vconsole.keymap=us crashkernel=auto
   vconsole.font=latarcyrheb-sun16 rhgb quiet intel_idle.max_cstate=0
   processor.max_cstate=1"

   grub2-mkconfig -o /boot/grub2/grub.cfg
   ```

6. Disable firewall, `WARN (N0010)`

   ```
   iptables -F

   iptables -t nat -F

   iptables -t mangle -F

   iptables -X

   systemctl mask firewalld

   systemctl disable firewalld

   systemctl stop firewalld
   ```

7. Change SELinux, `FAIL (S0081)`

   ```
   vi /etc/selinux/config
   ```

   ```
   SELINUX=permissive
   ```

8. Configure the BIOS for maximum performance.

   System Configuration > BIOS/Platform Configuration (RBSU) > Power Management > HPE Power Profile > [Maximum Performance]

9. Reboot the system.
10. After the nodes reboot, verify that the limits have been increased.

    ```
    ulimit -a
    ```

References

- Configuring the HPE Proliant DL380 Gen9 24-SFF CTO Server as a Vertica Node

  `https://my.vertica.com/kb/Configuring-the-HPE-Proliant-DL380-Gen9-24-SFF-CTO-Server-as-a/Content/Hardware/Configuring-the-HPE-Proliant-DL380-Gen9-24-SFF-CTO-Server-as-a.htm`

- Tuning Linux Dirty Data Parameters for Vertica

  `https://my.vertica.com/kb/Tuning-Linux-Dirty-Data-Parameters-for-Vertica/Content/BestPractices/Tuning-Linux-Dirty-Data-Parameters-for-Vertica.htm`

The following addresses problems that may be encountered during the Vertica installaton:

- Check for Swappiness

  S0112: `https://my.vertica.com/docs/8.1.x/HTML/index.htm#cshid=S0112`
- I/O scheduling

  S0150: `https://my.vertica.com/docs/8.0.x/HTML/index.htm#cshid=S0150`
- Disk readahead

  S0020:`https://my.vertica.com/docs/8.0.x/HTML/index.htm#cshid=S0020`
- CPU frequency scaling

  S0141: `https://my.vertica.com/docs/8.0.x/HTML/index.htm#cshid=S0141`
- Firewall considerations

  N0010: `https://my.vertica.com/docs/8.0.x/HTML/index.htm#cshid=N0010`
- SELinux configuration

  S0081: `https://my.vertica.com/docs/8.0.x/HTML/index.htm#cshid=S0081`

Generating the SSH key pair

About

Generate a key pair on node 1 and then copy the public key to all nodes in the cluster, including node 1. This enables password-less SSH access from the node 1 server to all the other node servers in the cluster.

Procedure

1. On the node 1 server, run the `ssh-keygen` command.

   Example:

   ```
   ssh-keygen -q -t rsa
   ```
2. Copy the key from node 1 to all nodes, including node 1, using the node IP address.

   Example:

   ```
   ssh-copy-id -i ~/.ssh/id_rsa.pub root@11.111.111.111
   ```
   The system displays the key fingerprint and requests to authenticate with the node server.
3. Enter the required credentials of the nodes.

   The operation is successful when the system displays the following message:

   ```
   Number of key(s) added: 1
   ```
4. To verify that the key was successfully installed on the node, run the following command from node 1 to the target node to verify that it can successfully log into the node.

   ```
   ssh 'root@11.111.111.111'
   ```
5. Repeat steps 1 through 4 for all nodes.

Installing Vertica

About

Install no other application but Vertica on the Vertica server.

Procedure

1. On the Vertica cluster node 1 server, create a folder for the ArcSight Investigate Vertica Database Installer script.

   ```
   mkdir vertica-install-DIR
   ```
2. Copy the Investigate Vertica Database scripts.

   `arcsight-investigate-vertica-scripts.<hash>.tar.gz and arcsight-investigate-vertica-scripts.<hash>.tar.gz.md5` to `vertica-install-DIR`
3. Verify that the tarball matches the MD5 checksum.

```
cd vertica-install-DIR
md5sum arcsight-investigate-vertica-scripts.<hash>.tar.gz
cat arcsight-investigate-vertica-scripts.<hash>.tar.gz.md5
```

Both outputs should match.

4. Extract the tar file.

```
tar xvfz arcsight-investigate-vertica-scripts.<hash>.tar.gz
```

5. Edit the `config/vertica_user.properties` file.

The `hosts` and `license` properties must be updated.

Property	Description
hosts	A comma separated list of the Investigate Vertica Database servers in IPv4 format (1.1.1.1)
license	Download the license file from the Software Entitlements portal. Place the license file on your file system, and then point to this file from license parameter.
db_retention_day	Use for the data retention policy. See "Managing the data retention policy on the Vertica cluster" on page 43.

6. Install Vertica.

```
./vertica_installer install
```

You will be prompted to set up two users, a database administrator and an Investigate search user.

See **Reference** for additional Vertica Installer options.

After installation completes, safeguard your database admin credentials (example: `config/vertica_user.properties`).

7. Create the schema.

```
./vertica_installer create-schema
```

8. Create a scheduler.

```
./kafka_scheduler create <EB Node 1 IP>:9092,<EB Node 2 IP>:9092,<EB Node 3 IP>:9092
```

See the "Reference" section for additional Kafka Scheduler options.

9. Check the Kafka Scheduler.
   a. To check the Kafka Scheduler status, run:

```
./kafka_scheduler status.
```

   b. To check event-copy progress, run:

```
./kafka_scheduler events
```

c. To check Kafka Scheduler messages, run:

```
./kafka_scheduler messages
```

Reference

- Additional Vertica Installer (`./vertica_installer <option>`) options:

Option	Description
install	Installs Vertica on hosts specified in: vertica_user.properties Avoid using local loopback (localhost, 127.0.0.1, etc) if cluster needs to be constructed. Example: vertica_installer install
create-schema	Create Vertica database with dba_password and for the hosts specified in: vertica_credentials.properties . Example: vertica_installer create-schema
delete-schema	Delete Vertica database with dba_password and for the hosts specified in: vertica_credentials.properties . Example: vertica_installer delete-schema
start-db	Starts Vertica database with dba_password both specified in: vertica_credentials.properties Example: vertica_installer start-db
status	prints Vertica-cluster status Example: ./vertica_installer status

- Additional Kafka Scheduler (`./kafka_scheduler <option>`) options:

Option	Description
create	Creates a new scheduler for given Kafka broker list. Example: `./kafka_scheduler create 192.168.1.1:9092,192.168.1.2:9092` ./kafka_scheduler create <EB Worker Node 1 IP>:9092,<EB Worker Node 2 IP>:9092,<EB Worker Node 3 IP>:9092 number_of_partitions
update	Updates the scheduler. Example: ./kafka_scheduler update 192.168.1.1:9092,192.168.1.2:9092
start	Starts scheduler for all Kafka instances registered after performing a stop operation first. Begins copying data from all registered Kafka brokers. Example: ./kafka_scheduler start
stop	Stops copying data from all registered Kafka brokers. Example: ./kafka_scheduler stop
delete	Deletes all registered Kafka instances and meta data from the scheduler. After doing this, immediately run the kafka_scheduler create command. Example: ./kafka_scheduler delete

Option	Description
status	Print information and log status for a running or stopped Scheduler, including the following: ○ Current Kafka cluster assigned to the scheduler ○ Name and Vertica host where the active scheduler is running ○ Name, Vertica host, and process ID of every running scheduler (active or backup) Example: `./kafka_scheduler status`
events	Prints event-copy progress for the scheduler. This includes the following: ○ Current event and reject event counts ○ Last (N*2) microbatch statements, where N is the number of partitions of the source topic
messages	Prints the last ten log messages reported by the scheduler. You are prompted to review additional logs in the scheduler log file.

See Also

Troubleshooting

Chapter 5: Install ArcSight Investigate

ArcSight Installer is used to install ArcSight Investigate and Event Broker. The installer also configures firewall settings during setup (in case `firewalld.service` is up and running) on both the Kubernetes master and worker nodes.

Before deploying Investigate, you must download the docker images. There are two options available to get these images. You choose which approach meets your specific needs.

- Online mechanism: Pull the images from docker.com using the download_images script in packaged in ArcSight Installer.
- Offline mechanism: Downloading a tar file from the Investigate software download site.
- If you have chosen a multi-master deployment for ArcSight Event Broker, we recommend that you install Investigateon a worker node. It is your option whether to install in on a worker node dedicated to Investigate or on a worker node that also has Event Brokerpods running.

Multi-master installation is not supported.

Labeling nodes

About

- The deployment typically consists of three Kubernetes master nodes and three Kubernetes worker nodes. Also supported is a configuration with a single Kubernetes master node, and three Kubernetes worker nodes. In this case, the ArcSight Event Broker should be installed on the three worker nodes and ArcSight Investigate should be installed on the master node.
- You can add additional worker nodes to extend the Kafka cluster nodes (see the *ArcSight Event Broker Deployment Guide*.
- Once you add the new worker nodes, labels can be used to assign specific pods to them, like with Kafka.

Procedure

SSH to the master node, label all nodes, and then confirm that the label was applied.

```
# kubectl label --overwrite node <node_ip> investigate=yes

# kubectl get nodes -L investigate

NAME             STATUS    AGE       VERSION   INVESTIGATE

<node_ip>   Ready     2d        v1.6.1    <none>

<node_ip>   Ready     2d        v1.6.1    yes
```

```
<node_ip>   Ready     2d        v1.6.1    <none>
<node_ip>    Ready     2d        v1.6.1    <none>
```

Obtaining Investigate images online for the Installer

Procedure

1. Download ArcSight Investigate images.

   ```
   cd /opt/arcsight/kubernetes/scripts
   ./downloadimages.sh --suite investigate --registry docker --org arcsightsecurity
   ```
2. Pick the 2.10 version.
3. Upload the images to the local Docker registry.

   ```
   ./uploadimages.sh --suite investigate
   ```
4. If not already deployed, obtain Event Broker images.

   See the *ArcSight Event Broker Deployment Guide*.

Obtaining Investigate images offline for the Installer

About

If your environment does not permit internet access, you can download the ArcSight Investigate images offline.

Procedure

1. Download the images tar file (`arcsight-investigate-*.tar`) from the ArcSight Investigate software entitlement site.
2. Verifying the download.

   Micro Focus provides a digital public key to enable you to verify that signed software you download from the software entitlement site is indeed from Micro Focus and has not been manipulated in any way by a third party. Visit the following site for information and instructions:

   `https://h20392.www2.hpe.com/portal/swdepot/displayProductInfo.do?productNumber=HPLinuxCode Signing`
3. Place `arcsight-investigate-*.tar` in `master:<offline install directory>`.

   The `<offline install directory>` is a local directory where you can conveniently access the tar file.
4. Upload the Investigate images to make them available to the Installer.

```
cd <offline install directory>
tar xvf arcsight-investigate-*.tar
```

All Investigate related images will be stored in the `./investigate` directory.

```
cd/opt/arcsight/kubernetes/scripts
./uploadimages.sh --suite investigate --dir <offline install
directory>/investigate
```

ArcSight Installer tasks

From the ArcSight Installer (UI page), you can check the status of the master and worker nodes, deploy the Investigate images, and configure Investigate .

Deploying Investigate images

About

The Investigate deployment option is located in the Deployment page of the ArcSight Installer, having an initial status of **OFF**.

Procedure

Location: ArcSight Installer

1. Login to the ArcSight Installer (UI page).

   `https://<master-FQDN>:5443`

2. In the the left navigation pane, click **the Deployment** link.
3. In the ArcSight Investigate row, click **the Deploy button**.
4. In the version dialog, select 2.10 and then click Deploy. You will see progress indicator in the ArcSight Investigate row under the Status header. You will see a temporary drop down notification appear when deployment has started.
5. You can check the pod status using either of the following approaches.
   - On the ArcSight Installer Deployment page, when the ellipses icon appears in the Details column, click that icon. A dialog will appear with a summary status of each pod.
   - Connect via ssh to a master kubernetes node, then run the command, `kubectl get pods --all-namespaces`. If the pods are healthy the status of each pods will be Running.

   NOTE: It may take a few minutes for all pods to start running.

If the Vertica database connection has not been configured, or is configured incorrectly, the hecules-search-* container will have an error status on the ArcSight Installer Deployment page and CrashLoopBackOff status at the command line.

6. Retrieve the search user credentials that you defined when installing Vertica.

   `config/vertica_user.properties`

   If Vertica is not installed, install the database and then retrieve the search-user credentials (see "Installing Vertica" on page 23).

7. From the left navigation, choose **Configuration > ArcSight Investigate > Vertica**.
8. Configure the InvestigateVertica database connection (see "Configuring the ArcSight Investigate Vertica database connection" on page 40).

   NOTE: Each time you deploy, make sure to reconfigure the Investigate Vertica database connection. The information does not persist when the application is undeployed.

Each time you change the Vertica database connection, the `hercules-search-*` container will restart. Check the status of the container using one of the following methods.

- On the ArcSight Installer Deployment page, when the ellipses icon appears in the Details column, click that icon. A dialog will appear with a summary status of each pod.
- Connect via ssh to one of a master kubernetes node, then run the command, `kubectl get pods --all-namespaces` . If the pods are healthy the status of each pods will be Running.

See also

"Undeploying Investigate" on the next page

Configuring Event Broker for ArcSight Investigate

About

Once you deploy Event Broker, you can then configure the Event Broker data pipeline for ArcSight Investigate from the ArcSight Installer.

Notes:

- Event Broker consumers need a signed certificate from the Event Broker to establish secure communication with Investigate (see the *ADP Event Broker Deployment Guide*).
- In the event of a planned redeployment of Event Broker without a restart of the cluster node systems, be sure to do a clean undeploy of event broker (see the *ADP Event Broker Deployment Guide*).

Procedure

Location: ArcSight Installer > left navigation > Configuration

1. Select **ArcSight Event Broker**.
2. Select **Replicas**.
3. Click **+** next to Transforming String Processor and click **Save**.
   The number will change from 0 to 1.

Undeploying Investigate

Procedure

Location: ArcSight Installer

1. Login to the ArcSight Installer (UI page).

   `https://<master-FQDN>:5443`

2. From the left navigation, click **Deployment**.
3. Click **Undeploy** for ArcSight Investigate 2.10.

See also

"Deploying Investigate images" on page 29

Chapter 6: Upgrade ArcSight Investigate

ArcSight supports an upgrade from ArcSight Investigate 2.01 to version 2.10. Upgrading involves updating the Vertica Installer and database, Investigate, Event Broker, and ArcSight Installer. See the *ArcSightEvent Broker Deployment Guide* to update the last two of these.

General upgrade requirements

- Make a configuration backup before upgrading to this release.

  Refer to the *ArcSight Investigate User's Guide* for the Investigate version you are currently running.
- You may need to upgrade your operating system to a supported version before upgrading Investigate.

  See "Supported operating systems " on page 14.
- Ensure ArcSight Installer is at version 1.40.

  If you have an older version, see the *ArcSight Event Broker 2.20 Release Notes* for upgrade details.
- Ensure ArcSightEvent Broker is at version 2.20.

  If you have an older version, see the *ArcSight Event Broker 2.20 Release Notes* for upgrade details.

Back Up Existing Lookup Lists

About

If you are using a lookup list and do not have the original CSV file, then backup the lookup list before migrating to Investigate 2.10. This procedure is required to migrate the IP and MAC addresses from string to ByteArray datatype.

Procedure

1. Create a fieldset with the lookup list fields selected.
2. Remove all fields except the one to be used ~~on the LHS of~~ to the left of the `in list` operator in the search query.
3. In the search field, specify the field using the `in list` operator and then pick the lookup list.

   Example:

   `Source Address in list l1_srcAddr`
4. Add an inline filter for `Source Address`.

   This is the same field used ~~on the LHS of~~ to the left of the `in list` operator in search query.

Example:

```
Source Address <hide duplicates> <empty value>
```

5. Run the search and then export the results to the CSV file from the Events table.
6. After migrating to Investigate 2.10 and before uploading the CSV file of the lookup list backup, make the following changes to the CSV file:
   - Remove the field that is to the left of the query.

     In this example, `Source Address` is not part of the lookup list.
   - Edit the column headers which will be <lookup list name>_<lookup fieldname> to just be <lookup field name>

     ** Comment ** For the above bullet: Is this accurate? Should it read as follows... Edit the column headers <lookup list name>_<lookup fieldname> to be <lookup field name>.
7. You will re-upload the CSV files after migrating all Investigate deployment components to update the Lookup List datatypes.

Upgrading the Vertica Installer

About

Upgrading the Vertica Installer gives you the ability to upgrade to the latest version of ArcSight Investigate.

Procedure

1. Download the Investigate 2.10 Vertica installer file (`.tar.gz`) and `md5` file to a temporary location on the primary Vertica server.

   The primary Vertica server is the same system where you ran the Investigate Vertica installation scripts when you set up the cluster for the first time.
2. Check the `md5sum` of the `tar.gz` and cat the `md5` file to ensure they are the same.
3. Untar the Investigate 2.10 Vertica installer `tar.gz` file.
4. Run the `investigate_upgrade` script as root user.

   This script does the following:
   - Assumes existing Investigate Vertica utilities were installed under `/root/install-vertica`, the default location. It will fail if that location does not exist.
   - Adds the data retention script and updates the `vertica.properties` file with new configuration parameters, but does not enable the data retention.

     To enable data retention, see "Managing the data retention policy on the Vertica cluster" on page 43.
5. Delete the directories where the Investigate 2.10 file was untarred.

Upgrading Vertica

About

Before upgrading Vertica, ensure that your Vertica cluster is stable, with no current error messages in `vertica.log`.

Ensure no other application but Vertica is installed on the Vertica server.

Procedure

1. On the Vertica cluster node 1 server, create a folder for the new ArcSight Investigate Vertica Database Installer script.

   ```
   mkdir new-vertica-install-DIR
   ```

2. Copy the Investigate Vertica install scripts.

   ```
   arcsight-investigate-vertica-scripts.<hash>.tar.gz and
   arcsightinvestigate-

   vertica-scripts.<hash>.tar.gz.md5 to new-vertica-install-DIR
   ```

3. Verify that the tarball matches the MD5 checksum.

   ```
   cd new-vertica-install-DIR

   md5sum arcsight-investigate-vertica-scripts.<hash>.tar.gz

   cat arcsight-investigate-vertica-scripts.<hash>.tar.gz.md5
   ```

   Both outputs should match.

4. Extract the tar file.

   ```
   tar xvfz arcsight-investigate-vertica-scripts.<hash>.tar.gz
   ```

   Note: The original directory for the ArcSight Investigate Vertica Database Installer script is `original-installer-script-DIR`.

5. Run `./investigate_upgrade -c upgrade-investigate`

   The following is typical output, if using `/opt/v1310` as the original install location:

   ```
   ********************* Start of Investigate Upgrade *****************

   Enter previous installed location (/root/install-vertica):/opt/v1310

   Property File is missing, File: /opt/v1310/config/vertica_user.properties
   ```

```
Property File is missing, File: /opt/v1310/config/vertica_
credentials.properties

Checking all Vertica nodes are UP

All Vertica nodes are UP

Current Investigate version is: INIT_VER

Investigate will be upgraded to 2.10.0

Upgrading script and config files.

Creating backup directory: /opt/v1310/oldVersion

Backing up: /opt/v1310/vertica.properties

Backing up: /opt/v1310/scripts

Backing up: /opt/v1310/kafka_scheduler

Backing up: /opt/v1310/vertica_installer

Backing up: /opt/v1310/data

Upgrading: /opt/v1310/investigate_upgrade

Upgrading: /opt/v1310/vertica.properties

Upgrading: /opt/v1310/vertica_upgrade.py

Upgrading: /opt/v1310/scripts

Upgrading: /opt/v1310/kafka_scheduler

Upgrading: /opt/v1310/arcsight-vertica-installer-2.10.0.60.tar.gz

Upgrading: /opt/v1310/vertica_installer

Upgrading: /opt/v1310/data

Upgrading: /opt/v1310/upgrade

Upgrading: /opt/v1310/vertica-upgrade.log

************ Investigate Upgraded Complete. Version is 2.10.0 ************
```

6. Run:

   ```
   cd /opt/v1310

   ./kafka_scheduler update
   ```

   The following is typical output:

```
Terminating all running scheduler processes for schema: [investigation_
scheduler]

scheduler instance(s) deleted for 192.214.137.73

scheduler: update microbatch for mbatch_192_214_136_164

Successfully updated kafka scheduler.

Terminating all running scheduler processes for schema: [investigation_
scheduler]

scheduler instance(s) deleted for 192.214.137.73

scheduler instance(s) added for 192.214.137.73
```

7. Results of the upgrade:
   - File changes
     - Two files split off from `vertica.properties`:
       - `/opt/v1310/config/vertica_user.properties`
       - `/opt/v1310/config/vertica_credentials.properties`

       Safeguard both files.

       The installation moves properties to the new files.

Old file	New file	Properties
vertica.property	config/vertica_user.properties	hosts=localhost license=/your/license/file/here.da
vertica.property	config/vertica_credentials.properties	~~dba_user=<change-me>~~ ~~dba_password=<change-me>~~ ~~search_user=<change-me>~~ ~~search_password=<change-me>~~
config/sched.properties	config/vertica_credentials.properties	~~ingest_user=ingest~~ ~~ingest_password=ingest~~

   - After upgrade, the string IP and MAC address columns will not receive any new data during ingestion. Existing values in those columns will be left as is. There are 18 columns in the `investigation.events` table that fall into this group.
   - The following are create and insert into the `investigation.version_metadata` table:
     - `schemaVersion='4.4.0'`
     - `installerVersion='2.10.0'`

- All old files (except config directory files from `/root/install-vertica/`) are replaced with new files from `/tmp/upgrade-vertica`.
- The 'target_columns' column in the table investigation_scheduler.stream_microbatches are updated to remove 18 non-Bin IP and MAC column names. Now, the list will only contain IP/MAC address column names that end with 'Bin'.

8. To check the Kafka Scheduler status, run:

   ```
   ./kafka_scheduler status.
   ```

9. Copy the line this line `#db_retention_days=90` from the original vertica.properties file and add it to the end of the config/vertica_user.properties file.
10. To check event-copy progress, run:

    ```
    ./kafka_scheduler events
    ```

11. To check Kafka Scheduler messages, run:

    ```
    ./kafka_scheduler messages
    ```

Upgrading Investigate

About

The ArcSight Investigate upgrade process supports only an offline upgrade.

Procedure

1. Download the ArcSight Investigate offline upgrade file (`.tar.gz`) from the Micro Focus Customer Support site at Micro Focus Software Support.

   Investigate documentation is not included in your download package. Download your documentation from the ArcSight Product Documentation Community on Protect 724.
2. Extract the offline upgrade file to:

   `/opt/arcsight/upgrade/investigate-2.10`.
3. Upload images for Investigate.

   ```
   cd /opt/arcsight/kubernetes/scripts
   ./uploadimages.sh --suite investigate --dir
   /opt/arcsight/upgrade/investigate-2.10
   ```

4. Login to the ArcSight Installer (UI page).

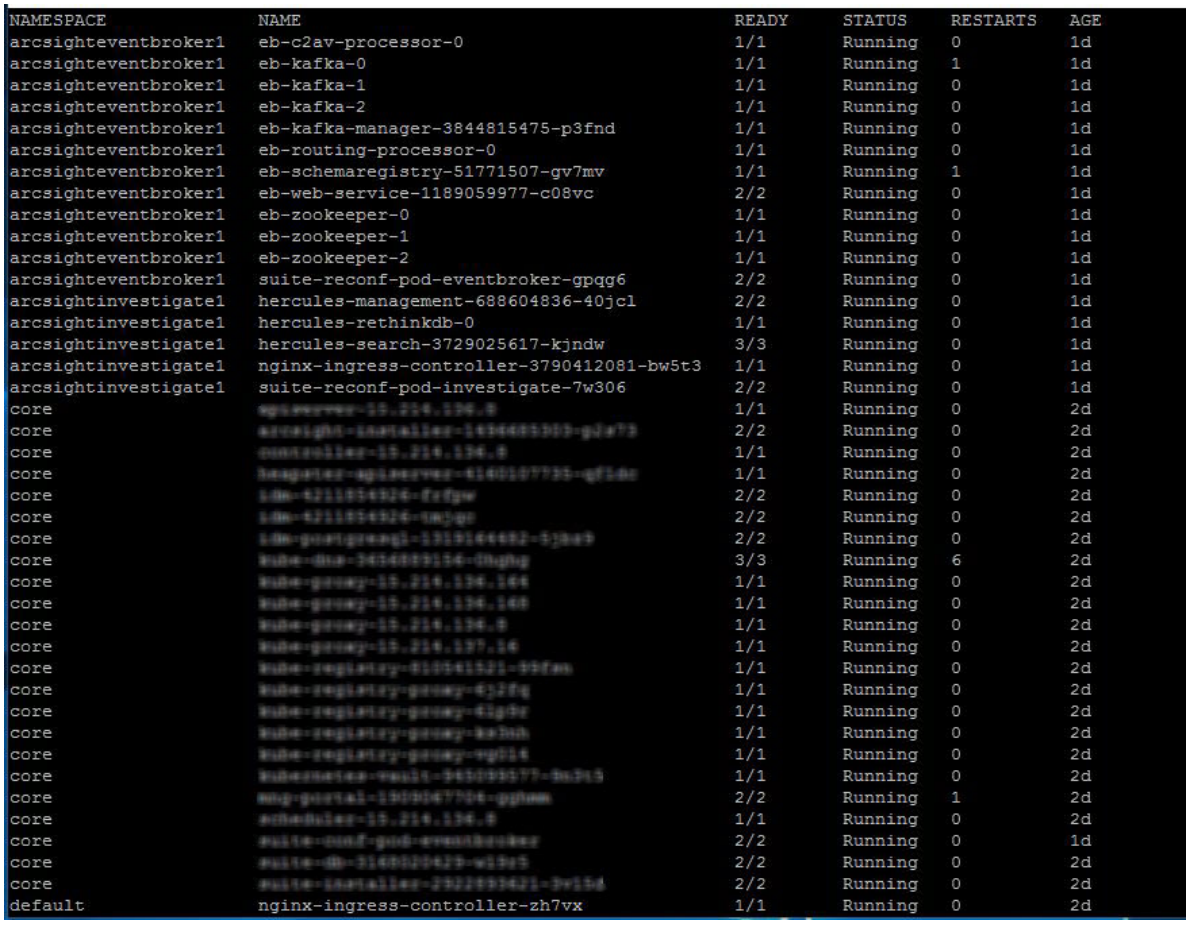
   `https://<master-FQDN>:5443`
5. From the ArcSight Installer (UI page), click **Node Management** and ensure that worker nodes are running and all have the status of READY.
6. From the Deployment page, click **Upgrade** for ArcSight Investigate.

7. Configure Event Broker for ArcSight Investigate.
   a. From the Configuration page, select **ArcSight Event Broker**.
   b. Select **Replicas**.
   c. Click **+** next to Transforming String Processor and click **Save**.
      The number will change from 0 to 1.
8. You can check the pod status using either of the following approaches.
   - On the ArcSight Installer Deployment page, when the ellipses icon appears in the Details column, click that icon. A dialog will appear with a summary status of each pod.
   - Connect via ssh to a master kubernetes node, then run the command, `kubectl get pods --all-namespaces`. If the pods are healthy the status of each pods will be Running.

   NOTE: It may take a few minutes for all pods to start running.

   When deployment is complete, all pods should be in the running state, as shown in the example below.

   Note: It may take a few minutes for all pods to start running.



NAMESPACE	NAME	READY	STATUS	RESTARTS	AGE
arcsighteventbroker1	eb-c2av-processor-0	1/1	Running	0	1d
arcsighteventbroker1	eb-kafka-0	1/1	Running	1	1d
arcsighteventbroker1	eb-kafka-1	1/1	Running	0	1d
arcsighteventbroker1	eb-kafka-2	1/1	Running	0	1d
arcsighteventbroker1	eb-kafka-manager-3844815475-p3fnd	1/1	Running	0	1d
arcsighteventbroker1	eb-routing-processor-0	1/1	Running	0	1d
arcsighteventbroker1	eb-schemaregistry-51771507-gv7mv	1/1	Running	1	1d
arcsighteventbroker1	eb-web-service-1189059977-c08vc	2/2	Running	0	1d
arcsighteventbroker1	eb-zookeeper-0	1/1	Running	0	1d
arcsighteventbroker1	eb-zookeeper-1	1/1	Running	0	1d
arcsighteventbroker1	eb-zookeeper-2	1/1	Running	0	1d
arcsighteventbroker1	suite-reconf-pod-eventbroker-gpqg6	2/2	Running	0	1d
arcsightinvestigate1	hercules-management-688604836-40jc1	2/2	Running	0	1d
arcsightinvestigate1	hercules-rethinkdb-0	1/1	Running	0	1d
arcsightinvestigate1	hercules-search-3729025617-kjndw	3/3	Running	0	1d
arcsightinvestigate1	nginx-ingress-controller-3790412081-bw5t3	1/1	Running	0	1d
arcsightinvestigate1	suite-reconf-pod-investigate-7w306	2/2	Running	0	1d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	2/2	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	2/2	Running	0	2d
core	[redacted]	2/2	Running	0	2d
core	[redacted]	2/2	Running	0	2d
core	[redacted]	3/3	Running	6	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	2/2	Running	1	2d
core	[redacted]	1/1	Running	0	2d
core	[redacted]	2/2	Running	0	1d
core	[redacted]	2/2	Running	0	2d
core	[redacted]	2/2	Running	0	2d
default	nginx-ingress-controller-zh7vx	1/1	Running	0	2d

See also

To configure the session and search settings, see "Configuring session and search settings in ArcSight Installer" on page 41.

Migrating Investigate search components

About

The final phase of the upgrade requires manual steps involving the lookup list, existing searches, and dashboard charts.

Procedure

1. Re-upload the CSV files for each Lookup list or restore the lookup list backup that you created in "Back Up Existing Lookup Lists " on page 32.
2. Re-execute existing searches.

   Existing searches will produce errors. To resolve this, change the time and re-execute the search. You can then go back to the original search time and re-execute the original search.
3. Refresh or recreate dashboard charts.

   Two different time ranges are supported on the dashboard, one is fixed and the other is custom, such as the last one minute. Only widgets having a custom time range can be refreshed.

   Upgrading to Investigate 2.10 will cause existing dashboard widgets to throw errors. To deal with this, do the following:

   a. Refresh all charts having a custom time range (**... > Refresh**).

   b. For chats having a fixed time range, choose **... > Create search** to navigate to the Search page where you can create a new search using the query from the dashboard chart.

      The search will also have the same chart added and the search will display existing data. You can re-run the search, or without re-running the search you can add the chart to the dashboard and delete the older chart from the dashboard.

Chapter 7: Configure ArcSight Investigate and components

Once you deploy ArcSight Investigate, you can then configure the product from the Configuration page of the Installer. After changing a product setting, Investigate restarts. Wait until restart completes before logging into Investigate.

Establishing the system admin

About

When you log in to ArcSight Investigate for the first time, you need to create the first user in the system. This user is assigned the system admin role.

Procedure

1. Open `https://<master-FQDN>` if Event Broker was deployed in single-master mode. Open https://<virtual-IP> if Event Broker was deployed in multi-master mode.
2. From the Welcome page, enter the name, email, and password information for the system admin and then click **Create System Admin**.
3. From the Login page, enter the credentials for the system admin.

Configuring the ArcSight Investigate Vertica database connection

Procedure

Location: ArcSight Installer

1. Login to the ArcSight Installer (UI page).

   `https://<master-FQDN>:5443`

   if Event Broker was deployed in single-master mode.

   `https://<virtual-IP>`

   if Event Broker was deployed in multi-master mode.
2. From the left navigation, choose **Configuration > ArcSight Investigate > Vertica tab**.
3. Enter the following information:

- Vertica host — <vertica-node1-IP>,<vertica-node2-IP>,<vertica-node3-IP>
- Vertica user name — See step 6 in "Installing Vertica" on page 23.
  This is the Investigate search-user name that you created during installation.
- Vertica database — Investigate. This was defined during schema creation and cannot be changed.
- Vertica password — See step 6 in "Installing Vertica" on page 23
  This is the Investigate search-user password.

4. Click **Save**.

Configuring the SMTP server

About

Configure access to your SMTP server in ArcSight Installer to enable users that you create in ArcSight Investigate to receive notification emails.

Procedure

Location: ArcSight Installer

1. Go to **Configuration > ArcSight Investigate** and then click the **User Management** tab.
2. In the **User Management** tab, enter the following information and then click **Save**:
   - SMTP Host
   - SMTP Port
   - SMTP User Name
   - SMTP Password
   - Sender Address

Configuring session and search settings in ArcSight Installer

About

You can configure the following properties:

- Session timeout

  When the user session ends, the user is redirected to the login screen in order to log in again. The default session timeout is 60 minutes.

- Search query timeout

  Search queries may take a long time and impact performance. You can put a limitation on the amount of time a search query runs. The default search query timeout is 60 minutes.

Procedure

Location: ArcSight Installer

1. Click **Configuration > ArcSight Investigate**.
2. From the **General** tab, do the following:
   - In the **Session timeout** field, enter the maximum time (in minutes) that you want a session to run.
   - In the **Search query timeout** field, enter the maximum time (in minutes) that you want a search query to run.
3. Click **Save**.

Configuring Vertica SSL

About

The ArcSight Installer contains the script, `/opt/arcsight/installer/k8s/master/cert-utils.sh` which provides a tool that enables you to generate a certificate signed by the root certificate authority used by Kubernetes and all modules.

Procedure

1. Connect to the master node (where installation will run) and run `./cert-utils.sh generate-certificate vertica => script` produce `vertica.key` and `vertica.crt`

   You can add `vertica` to your host name.
2. Copy `vertica.key` and `vertica.crt` to all Vertica nodes.

   It is also needed to copy there certificate of certificate authority (default `/opt/arcsight/kuberntes/ssl/ca.crt`)
3. On each node, run the `su - -c adminTools dbadmin => vertica` administration tool.

   Use the user-specified in Vertica configuration.
   a. From the **Main Menu** in the **Administration Tools**, select **Configuration Menu** and then click **OK**.
   b. From the **Configuration Menu**, select **Distribute Config Files** and then click **OK**.
   c. Select **SSL Keys** and then click **OK**.
   d. Select the database on which you want to distribute the files (the database from configuration), and then click **OK**.

e. Add the file locations for the `vertica.crt`, `vertica.key` and `ca.crt` (certificate authority certificate) files, and then click **OK** to distribute the files.

See Also

https://my.vertica.com/docs/8.0.x/HTML/index.htm#Authoring/Security/SSL/ConfiguringSSL.htm

Managing the data retention policy on the Vertica cluster

About

- The data retention policy is based on calendar days.
- You can purge data either in real time or by using a scheduled cron job.
- With data retention enabled, you can purge data in Vertica that is older than the retention period, which is from 1 to 366 days.
- The default data retention period is 90 days.

  This period includes the past 90 days, starting with the current day.
- Events saved in Vertica are based on the device timestamp (`deviceReceiptTime` field), which may not correlate with the current date.

  Example: If you run the script on 6/30/2018 and the `db_retention_days` property is set to 90, then data older than 04/01/2018 will be deleted.
- With data retention disabled, data can be retained for more than 366 days.

Caution: Backup data if needed. Purged data cannot be retrieved.

Procedure

1. Backup Vertica data as needed.

   See "Backing up the Vertica database" on page 52.
2. If data retention is disabled, run the following commands:

   ```
   # cd $Vertica_install_directory/scripts/

   # ./retention_policy_util.sh -h
   ```

   The following appears:

   ```
   ------------------------------------------------------------------------

   Retention policy is not enabled. Please uncomment (db_retention_days)
   property in (vertica_user.properties).

   ------------------------------------------------------------------------
   ```

3. Enable data retention.

   ```
   # vi ../vertica_user.properties
   ```

   Change `#db_retention_days=90` to:

   ```
   db_retention_days=90
   ```

   To ensure data retention is enabled:

   ```
   [root@n15-214-137-h74 scripts]# ./retention_policy_util.sh -h
   ```

   The following should appear:

   ```
   Configure Vertica retention policy by purging old data

   Usage: ./retention_policy_util.sh [options]

   Options:

   -h, --help              display this message

   -l, --list              list current configuration

   -u, --update            update current retention policy value in
   properties file (valid values: 1 - 366),(default: 90 days)

   -t, --time              list current day-count in Investigate database

   -e, --eval              evaluate your purge policy plan before running '-
   -purge' option

   -p, --purge             purge data older than retention period property.
   Hint: run '--eval' option first

   -s, --schedule          create cronjob schedule

   -d, --disable           disable cronjob schedule

   Examples:

   ./retention_policy_util.sh -u

   ./retention_policy_util.sh -u 100

   ./retention_policy_util.sh -p
   ```

4. Check how many days of data is in the Vertica database.

   ```
   # ./retention_policy_util.sh -t
   ```

   The following is what can typically appear:

   ```
   ------------------------------------------------------------------------
   ```

```
Investigate has 100 day(s) with time-range: [2017-10-26 - 2018-02-06].

--------------------------------------------------------------------------
```

Note: There are more than 100 calendar days between 2017-10-26 to 2018-02-06. The `-t` results shows there are only 100 event-days, meaning there are 100 days that have incoming events. The differences means there are certain calendar days without incoming events.

5. Check to determine how many days of data to retain.
   a. To view the current configuration:

```
# ./retention_policy_util.sh -l

--------------------------------------------------------------------------

Current retention value is set to: 90 day(s)

--------------------------------------------------------------------------
```

   b. Specify the data retention period, if not using the default.

```
# ./retention_policy_util.sh -u

--------------------------------------------------------------------------

Current retention value is set to: 90 day(s)

--------------------------------------------------------------------------

# ./retention_policy_util.sh -u 99

--------------------------------------------------------------------------

Current retention value is set to: 99 day(s)

--------------------------------------------------------------------------

# ./retention_policy_util.sh -l

--------------------------------------------------------------------------

Current retention value is set to: 99 day(s)

--------------------------------------------------------------------------

Note:  (90 day default has been change to 99 days

# tail ../vertica_user.properties

...

## Please, uncomment to enable this feature
```

```
db_retention_days=99
```

c. Evaluate your purge results based on the current retention policy plan.

```
# ./retention_policy_util.sh -e
```

This action purges no data, but shows for which days data can be purged. Use the evaluation option to see how many days of data can be purged based on the current retention policy.

Note: Always check the purge results before purging data.

The following is what can typically appear:

```
**************************************************************************

***** No data will be purged. This is only evaluation for your
retention policy *****

**************************************************************************

Will purge time range : [ 2017-10-26 - 2017-10-31 ].

Will purge day 1, (2017-10-26)

Will purge day 2, (2017-10-27)

Will purge day 3, (2017-10-28)

Will purge day 4, (2017-10-29)

Will purge day 5, (2017-10-31)

***** done *****
```

Note: If the calendar day for each purge day is not consecutive, then there are no events for that particular calendar day. In this example, there are no events for 2017-10-30.

6. Purge Vertica data.

- To purge data using a scheduled cron job:

```
# ./retention_policy_util.sh -s
```

The following appears:

```
Cronjob has been created: (59 23 * * * /root/install-
vertica/scripts/retention_policy_util.sh -p &>> /root/install-
vertica/slogfile)
```

- To verify the cron job:

```
# crontab -l | grep retention
```

The following appears:

```
59 23 * * * /root/install-vertica/scripts/retention_policy_util.sh -p
&>> /root/install-vertica/slogfile
```

- To disable scheduled cron job:

  ```
  # ./retention_policy_util.sh -d
  ```

  The following appears:

  ```
  Do you really want to remove current schedule?

  1) Yes

  2) No

  #? 1

  Cronjob has been removed: (59 23 * * * /root/install-
  vertica/scripts/retention_policy_util.sh -p &>> /root/install-
  vertica/slogfile)

  # crontab -l

  Line: "59 23 * * * /root/install-vertica/scripts/retention_policy_
  util.sh -p &>> /root/install-vertica/slogfile" is gone
  ```

- To purge data in real time:

  ```
  # ./retention_policy_util.sh -p
  ```

  The following appears:

  ```
  ******************************************************

  ***** Purging event-data based on retention policy  *****

  ******************************************************

  Purging day 1, (2017-10-26), data partition attempt result: Partition
  dropped

  Purging day 2, (2017-10-27), data partition attempt result: Partition

  dropped

  Purging day 3, (2017-10-28), data partition attempt result: Partition

  dropped

  Purging day 4, (2017-10-29), data partition attempt result: Partition

  dropped
  ```

```
Purging day 5, (2017-10-31), data partition attempt result: Partition
dropped
***** done *****
```

See Also

"Installing Vertica" on page 23

Chapter 8: Uninstalling ArcSight Investigate

About

Uninstalling ArcSight Investigate requires two basic steps:

- Uninstall Kubernetes.
- Uninstall the Vertica database.

Procedure

1. Uninstall Kubernetes.
   a. Run the following command on all the worker nodes and on the master node and then reboot:

      ```
      /opt/arcsight/kubernetes/uninstall.sh
      yes
      yes
      ```

   b. After the server reboots, run the following command on the master node.

      ```
      rm -rf /root/.kube
      ```

      If `/root/.docker` exists, run `rm -rf /root/.docker`

   c. On the worker node(s), run:

      ```
      rm -rf /root/.kube /root/arcsight-installer-worker
      ```

      Note: If you want to delete all your data as well, run the following command on the master node.

      ```
      rm -rf /opt/arcsight /opt/kubernetes
      ```

      Run the following command on the worker nodes:

      ```
      rm -rf /opt/arcsight
      ```

2. Uninstall the Vertica database.

   Run the following on all Vertica nodes:

   a. Stop the process.

      ```
      pkill -9 -f vertica #stop vertica process
      ```

   b. Remove the package.

      ```
      rpm -e vertica-8.1.1-3.x86_64 #delete vertica package
      ```

   c. If you want to delete the configuration file used with your installation, you can choose to delete the `/opt/vertica/` directory and all sub-directories using this command:

      ```
      rm –rf /opt/vertica #delete Vertica data
      ```

      See the *Vertica Analytics Platform Version 8.1.x Documentation*.

Chapter 9: Backup and restore

The backup and restore of Investigate involves the following:

- Backup and restore of the Vertica database
- Backup and restore of Investigate user management and search datastores.

You can restore any of this data as necessary.

Backup Vertica

When to perform a backup

- Prior to a Vertica upgrade
- Prior to adding or removing a Vertica node
- After adding or removing a Vertica node
- After recovering from a crash
- Routinely

Vertica backup requirements

General requirements

- Extra storage

  While a backup is taking place, the backup process can consume additional storage. The amount of space consumed depends on the size of your catalog and any objects that you drop during the backup. The backup process releases this storage once the backup is complete.
- Vertica-to-Vertica version

  Backups can only be restored to the same version of Vertica from which the original backup came. For instance, you cannot backup Vertica 8.01 and restore it to a Vertica 8.10.
- Stop ingesting events

  Ingesting events into the database (through the Investigate Scheduler) during backup may result in the most recently ingested events not being backed up. In order to ensure that all events are backed up, you should stop ingestion prior to starting the backup.
- Backup host

  For best network performance, each Vertica node should have its own backup host.

- Directory usage

  Use one directory on each Vertica node to store successive backups.

Backup locations

Vertica supports the following locations where backups can be saved:

- Local folder on the Vertica node
- Remote server

Backup host file system

Backups can only be performed on the following file system types:

- ext3
- ext4
- NFS

Required storage

It is recommended that each backup host have space for at least twice the database node footprint size. Consider your long-term backup storage needs.

If you are using a single backup location, the estimated storage space needed for the Vertica cluster can be calculated through the following Vertica operation:

```
dbadmin=> select sum(used_bytes) as total_used_bytes from v_monitor.storage_containers;
 total_used_bytes
------------------
       5717700329
(1 row)
```

If you are using multiple backup locations, one per node, then the storage needed by each node can be calculated through the following Vertica operation:

```
dbadmin=> select node_name, sum(used_bytes) as total_used_bytes from v_
monitor.storage_containers group by node_name;
        node_name       |   total_used_bytes
------------------------+--------------------
 v_investigate_node0002 | 1906279083
 v_investigate_node0003 | 1905384292
 v_investigate_node0001 | 1906036954
```

```
(3 rows)
```

Backup host preparation

- Multiple hosts can be used to backup a Vertica cluster. Each host must be prepared prior to backup.
- Remote backup hosts must have SSH access and password-less SSH setup from Vertica node 1 in order for the database administrator to access the hosts (see "Setting up password-less SSH" below).
- If one host is the backup destination for multiple Vertica nodes, then increase the maximum SSH connections on the backup host by increasing the `MaxStartups` parameter in `/etc/ssh/sshd_config`. The `MaxStartups` number should be greater than the number of nodes in the Vertica cluster.

Setting up password-less SSH

1. Login to the backup server
2. Create user `$db_admin`

   `$db_admin` is the administrator for the Vertica cluster.
3. Ensure `$db_admin` has write permission to the dedicated directory where the backup will be stored.
4. Login to Vertica node 1 as root.
5. Become the Vertica database admin.

   ```
   # su -l $db_admin
   ```
6. Setup password-less SSH for all backup servers.

   ```
   # ssh-keygen -t rsa
   ```

   ```
   # ssh-copy-id -i ~/.ssh/id_rsa.pub $db_admin@$back_up_server_ip
   ```

Backing up the Vertica database

About

Backup is performed by the database admin (`$db_admin`).

Procedure

1. Login to Vertica cluster node 1 as root.
2. Generate a backup configuration file.

   ```
   # su -l $db_admin
   ```

   ```
   # /opt/vertica/bin/vbr --setupconfig
   ```

General configuration options:

- Restore Points — Default is 52, assuming a weekly backup for one year.

  Multiple restore points gives you the option to recover from one of several backups. For example, if you specify 3, you have 1 current backup, and 3 backup archives. Vertica stores the value you enter as the `restorePointLimit` parameter in the `vbr` configuration file.

- Password Save — The backup configuration can optionally save the database admin password to avoid prompting in the future.

- Advanced Options — These options allow additional security measures. It is recommended that the default options are used.

Per node configuration options:

- Backup Host Name — For each Vertica node, you will be prompted to specify the backup host. This host can be either the local machine or a remote host.

- Backup Directory — For each backup host, the directory where the backup will be stored must be specified

Example:

```
Warning: This setup tool is deprecated, and will be removed in a future
version. Please use config file samples we provide in
/opt/vertica/share/vbr/example_configs instead of using this tool.

Snapshot name (backup_snapshot): vertica_backup

Number of restore points (1): 52

Specify objects (no default):

Object restore mode (coexist, createOrReplace or create)
(createOrReplace): createOrReplace

Vertica user name (dbadmin): $db_admin

Save password to avoid runtime prompt? (n) [y/n]: n

Node v_investigate_node0001

Backup host name (no default): [BACKUP HOST 1 IP]

Backup directory (no default): /opt/vertica/backup1

Node v_investigate_node0002

Backup host name (no default): [BACKUP HOST 2 IP]

Backup directory (no default): /opt/vertica/backup2

Node v_investigate_node0003
```

```
Backup host name (no default): [BACKUP HOST 3 IP]

Backup directory (no default): /opt/vertica/backup3

Change advanced settings? (n) [y/n]: n

Config file name (vertica_backup.ini):

Saved vbr config to vertica_backup.ini.
```

The `vertica_backup.ini` file is created in `/home/$db_admin`.

> **Note:** The config file is needed for all future backup and restore operations. Save it in a safe place.

```
# cat ./vertica_backup.ini

[Misc]

snapshotName = vertica_backup

restorePointLimit = 52

objectRestoreMode = createOrReplace

[Database]

dbName = investigate

dbUser = analyst

dbPromptForPassword = True

[Transmission]

[Mapping]

v_investigate_node0001 = [BACKUP HOST 1 IP]:/opt/vertica/backup1

v_investigate_node0002 = [BACKUP HOST 2 IP]:/opt/vertica/backup2

v_investigate_node0003 = [BACKUP HOST 3 IP]:/opt/vertica/backup3
```

3. Initialize the backup locations.

```
# /opt/vertica/bin/vbr --task init --config-file vertica_backup.ini

Initializing backup locations.

Backup locations initialized.
```

4. Stop the Investigate scheduler.

```
# exit
```

```
# cd /root/install-vertica

./kafka_scheduler stop
```

Stopping the Investigate Scheduler from writing data to the Vertica database ensures that you do not lose events during backup.

5. Backup Vertica data.

```
# su -l $db_admin

# /opt/vertica/bin/vbr --task backup --config-file vertica_backup.ini

Starting backup of database investigate.

Participating nodes: v_investigate_node0001.

Enter vertica password:

Snapshotting database.

Snapshot complete.

Approximate bytes to copy: 270383427 of 270383427 total.

[==================================================] 100%

Copying backup metadata.

Finalizing backup.

Backup complete!
```

6. Verify that the backup files were written to the backup locations.

```
# ssh [BACKUP HOST 1 IP] ls /opt/vertica/backup1

backup_manifest

Objects

Snapshots

# ssh [BACKUP HOST 2 IP] ls /opt/vertica/backup2

backup_manifest

Objects

Snapshots

# ssh [BACKUP HOST 3 IP] ls /opt/vertica/backup3

backup_manifest
```

```
Objects

Snapshots
```

Backing up Vertica incrementally

About

Incremental backups use the same setup as a full backup and just back up what changed from the previous full backup. When a full backup is executed using the same configuration file, subsequent backups are incremental. When you initiate an incremental backup, the `vbr` tool displays a backup size that is a portion of the total backup size. This portion represents the delta changes that will be backed up by the incremental backup.

Procedure

Run the following commands:

```
# /opt/vertica/bin/vbr --task backup --config-file vertica_backup.ini

Starting backup of database investigate.

Participating nodes: v_investigate_node0001.

Snapshotting database.

Snapshot complete.

Approximate bytes to copy: 2680 of 270455624 total.

[==================================================] 100%

Copying backup metadata.

Finalizing backup.

Backup complete!
```

Verifying the integrity of the Vertica database backup

About

The `full-check` option is used to verify the integrity of the Vertica database backup snapshots, which reports the following:

- Incomplete restore points
- Damaged restore points

- Missing backup files
- Unreferenced files

Note: Backup files generated by the interrupted process will remain in the backup location and subsequent backups will resume where the interrupted backup left off.

Backup operations are atomic, so interrupted backup operations will not affect previous backup files.

Procedure

- Run the following command:

```
# /opt/vertica/bin/vbr --task full-check --config-file vertica_backup.ini

Checking backup consistency.

List all snapshots in backup location:

Snapshot name and restore point: backup_snapshot_20180116_172347, nodes:
['v_investigate_node0001'].

Snapshot name and restore point: backup_snapshot_20180116_172253, nodes:
['v_investigate_node0001'].

Snapshot name and restore point: backup_snapshot_20180116_172236, nodes:
['v_investigate_node0001'].

Snapshot name and restore point: backup_snapshot_20180116_172310, nodes:
['v_investigate_node0001'].

Snapshot name and restore point: backup_snapshot_20180116_172158, nodes:
['v_investigate_node0001'].

Snapshots that have missing objects(hint: use 'vbr --task remove' to delete
these snapshots):

Backup locations have 0 unreferenced objects

Backup locations have 0 missing objects

Backup consistency check complete.
```

Manage existing backups

Viewing available backups

- Run the following command:

```
# /opt/vertica/bin/vbr --task listbackup --config-file vertica_backup.ini

backup                          backup_type  epoch   objects   nodes
(hosts) file_system_type

vertica_backup_20180104_142326   full         29                v_
investigate_node0001(10.12.57.27)   [Linux]
```

  The backup name is comprised of the snapshot name and backup timestamp.

  Example:

  snapshot name: `vertica_backup`

  timestamp: `20180104_142326`

Deleting a backup

About

Use only the `vbr` tool to delete a backup.

Procedure

Run the following commands:

```
# /opt/vertica/bin/vbr --task remove --config-file /backup/vertica_backup.ini
--archive 20180104_142326

# 20180104_142326 is the backup timestamp

Removing restore points: 20180104_142326

Remove complete!


# /opt/vertica/bin/vbr --task listbackup --config-file vertica_backup.ini

backup                backup_type   epoch   objects   nodes(hosts)
              file_system_type
```

Restore Vertica data

Vertica restoration requirements

- Backups can only be restored to the same version of Vertica that the original backup came from. For instance, you can not backup Vertica 8.01 and restore it to a Vertica 8.10
- Restore to a cluster that is identical to the cluster from which the backup originated.

Prior to restoring a Vertica cluster, the cluster must meet the following conditions:

- Target database is already created and can be empty
- Target database name matches the backup database name
- Target database is stopped
- All Vertica nodes in the target cluster are up
- All Vertica nodes in the target cluster have identical names to the original backup

Restoring the Vertica database

About

Data restorations is performed by the database admin.

Procedure

1. Rebuild a Vertica cluster identical to the original cluster.
2. Login to Vertica node 1.

   ```
   # su -l $db_admin
   ```

   `$db_admin` has password-less SSH to the `$db_admin` of backup host (see "Setting up password-less SSH" on page 52).
3. Copy `vertica_backup.ini` to `/home/$db_admin`.
4. Stop the database.

   ```
   # /opt/vertica/bin/admintools -t stop_db -d investigate -p *****

   Connecting to database

   Issuing shutdown command to database

   Database investigate stopped successfully
   ```
5. Restore the backup data.

```
# /opt/vertica/bin/vbr --task restore --config-file vertica_backup.ini

Starting full restore of database investigate.

Participating nodes: v_investigate_node0001, v_investigate_node0002, v_
investigate_node0003.

Restoring from restore point: investigate_backup_20180110_010826

Determining what data to restore from backup.

[==================================================] 100%

Approximate bytes to copy: 2246248425 of 2246250258 total.

Syncing data from backup to cluster nodes.

[==================================================] 100%

Restoring catalog.

Restore complete!
```

6. Start the database.

```
# /opt/vertica/bin/admintools --task start_db -d investigate -p *****

Starting nodes:

v_investigate_node0001 (127.0.0.1)

Starting Vertica on all nodes. Please wait, databases with a large catalog
may take a while to initialize.

Node Status: v_investigate_node0001: (DOWN)

Node Status: v_investigate_node0001: (DOWN)

Node Status: v_investigate_node0001: (DOWN)

Node Status: v_investigate_node0001: (DOWN)

Node Status: v_investigate_node0001: (UP)

Database investigate started successfully
```

7. Start the Investigate Scheduler.

```
# exit

# cd /root/install-vertica

# ./kafka_scheduler start
```

Backing up Investigate management and search datastores

About

We recommend that you identify a back up location that is not under the `/opt/arcsight` directory. Identify a local folder on the system or a remote location.

This procedure uses the `/opt/investigate/backup` directory as an example.

Prerequisite

Ensure that the Investigate management and rethinkdb pods are off.

- To verify whether the pods are off:

  ```
  # kubectl get pods -n arcsightinvestigate1
  ```

  If the pod `hercules-rethinkdb-0` or the pod `hercules-management` appear in the least, then these pods are on.
- To deactivate the `rethinkdb` pod:

  ```
  # kubectl scale statefulsets hercules-rethinkdb -n arcsightinvestigate1 --replicas=0
  ```
- Verify that the pod is off by executing the following command and looking for the absence of the `hercules-rethinkdb` pod:

  ```
  # kubectl get pods -n arcsightinvestigate1
  ```
- To deactivate the management pod:

  ```
  # kubectl scale deployment  hercules-management -n arcsightinvestigate1 --replicas=0
  ```
- Verify that the pod is off by executing the following command and looking for the absence of the `hercules-management` pod:

  ```
  # kubectl get pods -n arcsightinvestigate1
  ```

Procedure

1. Logout of any open Investigate sessions.
2. SSH to the Kubernetes cluster master node 1.
3. Run the following commands:

   ```
   # cd /opt/arcsight/volumes/investigate/
   ```

```
# mkdir /opt/investigate/backup

# cp –R * /opt/investigate/backup

# diff -r -s /opt/investigate/backup/mgmt
/opt/arcsight/volumes/investigate/mgmt

# diff -r -s /opt/investigate/backup/search
/opt/arcsight/volumes/investigate/search
```

A message should state that all files are identical. If they are not identical, repeat the procedure.

4. To activate the rethinkdb pod:

```
# kubectl scale statefulsets hercules-rethinkdb -n arcsightinvestigate1 --
replicas=1
```

5. To verify that the pod is running, execute the following command and look for the hercules `rethinkdb` pod:

```
# kubectl get pods -n arcsightinvestigate1
```

6. To activate the management pod:
   a. The exact name of the `hercules-management-160458690` pod is enviornment dependent, and can be discovered through the command:

   ```
   # kubectl scale deployment hercules-management -n arcsightinvestigate1
   --replicas=1
   ```

   b. Verify that the pod is running by executing the following command and looking for the hercules management pod:

   ```
   # kubectl get pods -n arcsightinvestigate1
   ```

Restore management and search data

Restoring Investigate management and search datastores

About

When restoring the Investigate management and search datastores, retain the original directory structure under `/opt/arcsight/volumes/investigate/`.

The management datastore will be restored to the `/opt/arcsight/volumes/investigate/mgmt/db/` directory. The search datastore will be restored to the `/opt/arcsight/volumes/investigate/search` directory.

Prerequisite

Back up the Investigate management and search datastores to the `/opt/investigate/backup` directory on the Kubernetes master node. Also, ensure that the Investigate management and rethinkdb pods are off. See "Backing up Investigate management and search datastores" on page 61.

Procedure

1. Logout of any open Investigate sessions.
2. SSH to the Kubernetes master node and then run the following commands:

   ```
   # cd /opt/investigate/backup
   ```

   ```
   # cp –R search/* /opt/arcsight/volumes/investigate/search
   ```

   Reply `yes` to overwrite files and folders.

   ```
   # cd /opt/arcsight/volumes/investigate/mgmt/db/
   ```

   ```
   # rm - rf h2.lock.db
   ```

   ```
   # cp /opt/investigate/backup/mgmt/db/h2.mv.db   .
   ```

   Reply `yes` to overwrite files and folders.

   ```
   # diff -r -s /opt/arcsight/volumes/investigate/mgmt/db/h2.mv.db
   /opt/investigate/backup/mgmt/db/h2.mv.db
   ```

   ```
   # diff -r -s /opt/investigate/backup/search
   /opt/arcsight/volumes/investigate/search
   ```

   A message should state that all files are identical. If they are not identical, repeat the procedure.
3. Change the permission of the Investigate directory.

   ```
   # chown 1999:1999 -R /opt/arcsight/volumes/investigate/
   ```

   The directory structures should look similar to the following:

Potential issues during backup and restore

Vertica downtime exceeds the retention time for the Kafka cluster

If the Vertica cluster downtime exceeds the retention time for the Kafka cluster, there is a chance that the Vertica-stored Kafka offset is no longer present in the Event Broker cluster. In this case, the scheduler will not be able to consume new data.

Basic steps:

1. Determine the last Kafka offset read by the Scheduler.
2. Confirm the Scheduler is no longer copying data.
3. Reset the Scheduler.

Step 1: Determine the last Kafka offset read by the Scheduler

Each Vertica node in the cluster will copy data from one or more Kafka partitions. In order to see the source Kafka topic, as well as the last offset copied from each partition, run the Vertica SQL commands below. This query also identifies the number of partitions. Use that value as part of the limit clause in the second query below.

```
dbadmin=> select * from investigation_scheduler.stream_sources;
id  |      source      | cluster | partitions | enabled
----+------------------+---------+------------+---------
1   | eb-internal-avro |       1 |          1 | t
(1 row)
```

```
dbadmin => select frame_start, DISTINCT source_partition, end_offset from
investigation_scheduler.stream_microbatch_history order by frame_start desc
limit 1;
frame_start              | source_partition | end_offset
-------------------------+------------------+------------
2018-01-16 18:19:06.381  |                0 |         12
```

If the end_offset + 1 no longer exists in the Kafka cluster for the topic's source_partition, then it is likely that the scheduler will no longer be able to read from Kafka until it has been given a valid offset.

Step 2: Confirm the Scheduler is no longer copying data

About

You can confirm whether the scheduler is copying data or not by checking the status and examining the last copied offset in the microbatch status. If the offset number is not increasing, then the scheduler can no longer find the valid offset and needs to be reset.

Procedure

- Check the Scheduler offsets.

```
# ./kafka_scheduler status

...

'investigation_scheduler' scheduler last 10 microbatch status:

       frame_start       |   source_name    | start_offset | end_offset |
end_reason   | partition_bytes | partition_messages

------------------------+-----------------+----+----+--------------+-----
-----------+-------------------

2018-01-17 09:29:01.604 | eb-internal-avro |  9 |  9 | END_OF_STREAM |
         0 |                 0

2018-01-17 09:28:51.595 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-17 09:28:41.586 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-17 09:28:31.577 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-17 09:28:05.824 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-17 09:27:55.524 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-17 09:27:45.515 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-17 09:27:35.507 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-16 18:19:06.381 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

2018-01-16 18:18:56.374 | eb-internal-avro |           9 |          9 |
END_OF_STREAM |               0 |                 0

(10 rows)
```

Step 3: Resetting the Scheduler

About

The Event Broker cluster may continue to receive data. If the Vertica cluster down time exceeds the Event Broker data retention time, there is a chance that the Vertica offset will no longer be valid. The Scheduler must be reset. To reset the Investigate Scheduler, delete the Scheduler and then recreate it.

Procedure

- Execute the following commands:

```
# ./kafka_scheduler delete

Are you sure that you want to DELETE scheduler metadata (y/n)?y

Terminating all running scheduler processes for schema: [investigation_
scheduler]

scheduler instance(s) deleted for 192.214.138.94

bash: /root/install-vertica/kafka_scheduler.log: No such file or directory

scheduler instance(s) deleted for 192.214.138.95

bash: /root/install-vertica/kafka_scheduler.log: No such file or directory

scheduler instance(s) deleted for 192.214.138.96

db cleanup: delete scheduler metadata

# ./kafka_scheduler create
192.214.137.72:9092,192.214.137.71:9092,192.214.136.7:9092

create scheduler under: investigation_scheduler

scheduler: create target topic

scheduler: create cluster for
192.214.137.72:9092,192.214.137.71:9092,192.214.136.7:9092

scheduler: create source topic for
192.214.137.72:9092,192.214.137.71:9092,192.214.136.7:9092

scheduler: create microbatch for
192.214.137.72:9092,192.214.137.71:9092,192.214.136.7:9092

scheduler instance(s) added for 192.214.138.94

scheduler instance(s) added for 192.214.138.95

scheduler instance(s) added for 192.214.138.96
```

Troubleshooting

Vertica Scheudler throws the exception '[Vertica][VJDBC](5156) ERROR: Unavailable: initiator locks for query - Locking failure...'

This is an informational message that Investigate schema is not accessible at that moment in time. This is expected Kafka and Vertica behavior that happens periodically. Users should not be concerned about it.

Installing the ArcSight Installer Platform fails

Contact Technical Support.

Where to find the logs

To troubleshoot issues, capture the following logs. Logs are found under the pod number.

- zookeeper_container.log
- kafka_container.log
- schema-registry_container.log
- webservice_container.log

Pod starting order

After deploying Event Broker, pods are configured to start in the following order. Downstream pods will not start until the dependencies are met.

1. A quorum of zookeeper pods in the cluster must be up (2 of 3, or 3 of 5). Total number of zookeepers must be odd.
2. All Kafka pods must be up
3. Schema Registry pod must be up
4. Bootstrap Web Service, Kafka Manager
5. Transformation Stream Processor, Routing Stream Processor

SSL connection error

These are warnings that occur if there is a connection issue between Kafka and consumer or producer.

kubectl command is returning refused or time-out connection

If the `kubectl` command is returning refused or time-out connection, make sure the proxy is unset before repeating the command.

Vertica Scheduler unable to read events from Kafka

- New set up: Vertica Kafka scheduler: Check that Kafka scheduler is configured to communicate to Kafka port 39092.
- Working at first, but stopped working: Offset is not recognized: In this scenario, the kafka scheduler fails to recognize offset ids of messages that are in the topic. It can happen if the kafka scheduler unexpectedly stops reading from the topic, and then is restarted.
  Solution: execute the Kafka_scheduler delete command to delete the meta data. After doing this, immediately run the Kafka_scheduler create command to set up the scheduler.
- New set up: Check the network connection.
- New set up and existing set up: Check whether the broker is down.
- Existing set up: you did not configure all brokers that contain the topic the consumer connects to, and the brokers which are configured in that consumer are down.
- New set up: If you are encountering SSL connection related errors, check the steps that you used to import certificates to both Event Broker and consumers.

Appendix B: FAQs

Which pods in Kubernetes comprise the ArcSight Investigate deployment?

Hercules pods: management, proxy, rethinkdb, search

Related topic: ArcSight Investigate and Event Broker prerequisites

Can I use my existing Event Broker v1.0 with ArcSight Investigate + Vertica?

No. ArcSight Investigate requires Event Broker 2.0. You can migrate your data from Event Broker 1.0 using the Event Broker Data Migration utility. Check with ArcSight Support about the availability of this tool.

Related topic: *Event Broker Data Migration Tech Note*.

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 Deployment Guide (Investigate 2.10)

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 arc-doc@hpe.com.

We appreciate your feedback!