Micro Focus
ArcSight Platform

Technical Requirements for the ArcSight Platform

Wednesday, July 13, 2022

This Technical Requirements document describes the requirements and guidelines for the ArcSight Platform 22.1. The platform enables you to deploy a combination of security, user, and entity solutions into a single cluster within the Container Deployment Foundation (CDF) environment. The core services for this CDF environment, including the Dashboard and user management, are provided by a common layer called Fusion.

Customers running on platforms not provided in this document or with untested configurations will be supported until the point Micro Focus determines the root cause is the untested platform or configuration. According to the standard defect-handling policies, Micro Focus will prioritize and fix issues we can reproduce on the tested platforms.

Upgrade Paths

The 22.1 release of the ArcSight Platform supports upgrade of all CDF-based components from ArcSight Platform version 21.1 except for the ArcSight Database. Instead, you must remove the 21.1 version of the ArcSight Database and perform a fresh installation of the 22.1 version. For more information about the change to the database, see "ArcSight Database Has Become Smarter!" in the *Administrator's Guide to ArcSight Platform 22.1*.

Micro Focus does not recommend that ArcSight Intelligence customers attempt to upgrade without consulting with Micro Focus as upgrading at this time will result in loss of baseline data. We are aware that this is inconvenient for customers, and we are currently working on addressing this limitation in an future release.

Additional Documentation

The ArcSight Platform documentation library includes the following resources.

- *Release Notes for ArcSight Platform 22.1*, which provides an overview of the products deployed in this suite and their latest features or updates.
- *Administrator's Guide for ArcSight Platform 22.1*, which contains installation, user, and deployment guidance for the ArcSight software products and components that you deploy in the containerized platform.
- *User's Guide for Fusion in the ArcSight Platform*, which is embedded in the product to provide both context-sensitive Help and conceptual information.
- Product Support Lifecycle Policy, which provides information on product support policies.

Software Requirements

This section lists the software needed to install and run the ArcSight Platform.

Category	Operating System
Certified OS (minimal installation)	Red Hat Enterprise Linux 8.4 (x86, x64)
Supported OS (minimal installation)	**For CDF:** • Red Hat Enterprise Linux 8.2 (x86, x64) • Red Hat Enterprise Linux 7.9 (x86, x64) • Red Hat Enterprise Linux 7.8 (x86, x64) • CentOS 8.2 (x86, x64) • CentOS 7.9 (x86, x64) • CentOS 7.8 (x86, x64) **For Database, FIPS compliant versions of the following operating systems:** • Red Hat Enterprise Linux 8.2 (x86, x64) • CentOS 8.1 (x86, x64) • CentOS 8.2 (x86, x64)
File systems	One of the following: • EXT3 • EXT4 (recommended) • Logical Volume Manager (LVM) • XFS
Data Collection	Certified: SmartConnector 8.3 or later Supported: SmartConnector 8.2
Browser	• Google Chrome • Mozilla Firefox Browsers should not use a proxy to access Container Deployment Foundation (CDF) applications because this might result in inaccessible web pages.

Data Types and Connectors Supported by Intelligence

This section describes the data types and SmartConnectors/FlexConnector types Intelligence supports.

Data Types	Supported Smart Connectors
Access	SmartConnector for Microsoft Windows Event Log – Native Application and System Event Support SmartConnector for Microsoft Windows Event Log – Unified Application and System Event Support
Active Directory	SmartConnector for Microsoft Windows Event Log – Native Application and System Event Support
VPN	SmartConnector for Microsoft Network Policy Server File SmartConnector for Pulse Secure Pulse Connect Secure Syslog SmartConnector for Citrix NetScaler Syslog SmartConnector for Nortel Contivity Switch Syslog
Web Proxy	SmartConnector for Microsoft Forefront Threat Management Gateway File SmartConnector for Squid Web Proxy Server File SmartConnector for Blue Coat Proxy SG Multiple Server File
Repository	FlexConnector Type - ArcSight FlexConnector Regex File

Additional Considerations

Consider the following:

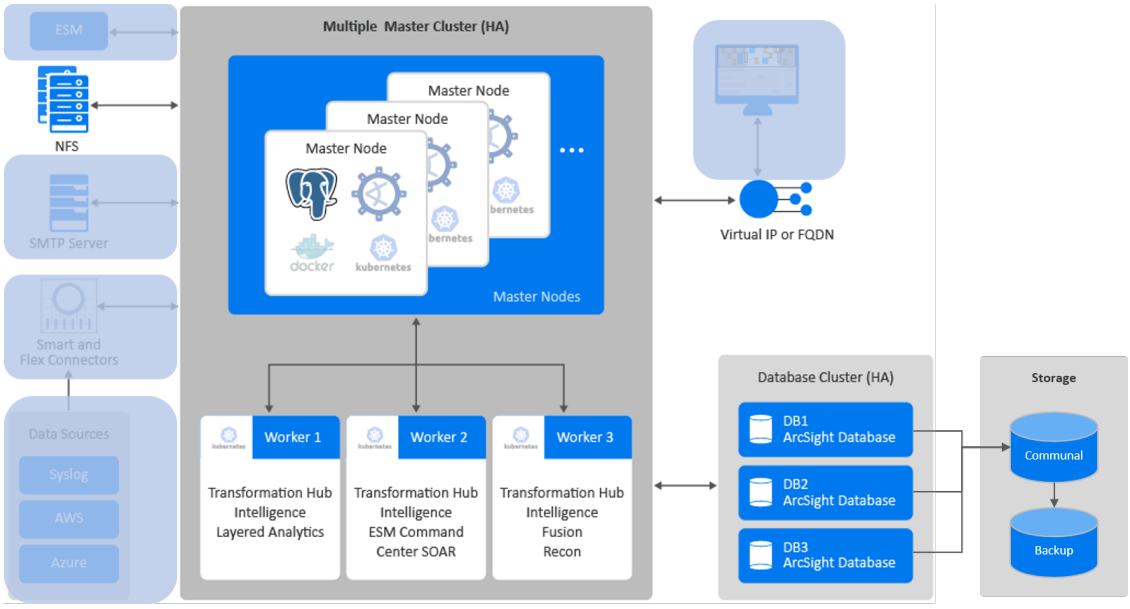
- A fuller set of SmartConnectors is supported for those sources that provide data of relevance to the Intelligence analytics models. Micro Focus might need to examine sample logs to optimize analysis of data from this broader set of sources.
- For supported data types, Intelligence provides support for new devices that provide data of relevance to the Intelligence analytics models. For more information, see "Adding Support for New Devices" in the *Administrators Guide for ArcSight Platform*.
- Intelligence supports the SmartConnectors listed. However, additional capabilities you might deploy, such as Recon, might support a wider set of SmartConnectors/FlexConnector types.

- Micro Focus advises against configuring event aggregation for data to be processed by ArcSight Intelligence. If you wish to use ArcSight Intelligence with aggregated events, contact Micro Focus Customer Support.

Hybrid Cloud Support

The ArcSight Platform can be deployed to a variety of locations, including on-premises and cloud. The CDF Infrastructure, capabilities deployed on it, and the ArcSight Database have been tested when deployed together to the same location. For these components, which are unshaded in the diagram below, the table below specifies the combination of location, Kubernetes service and ArcSight Database Communal Storage service that has been tested, whereas other combinations have not been tested.

Other related components can be deployed in a hybrid cloud manner, with some deployed to different locations than others. To understand the tested scenarios for these components, see the documentation related to that component.



Customers running on platforms not specified in this document or with untested configurations will be supported until the point Micro Focus determines the root cause is the untested platform or configuration. According to the standard defect-handling policies, Micro Focus will prioritize and fix issues we can reproduce on the tested platforms.

Location	Kubernetes	Communal Storage
Amazon Web Services (AWS)	Elastic Kubernetes Service (EKS)	AWS S3

Azure	Azure Kubernetes Service (AKS)	Azure Blob
On-premises	CDF embedded	"bring-your-own" On-premises S3-compatible storage

System Hardware Sizing and Tuning Guidelines

This section describes the system sizing and tuning for the ArcSight Platform and deployed capabilities that has been confirmed in our testing lab to maintain satisfactory performance of the system under a variety of workloads. These guidelines are based on dedicated resource allocations. In virtual environments, where there is a risk of over subscription of the physical hardware, ensure that the system meets this sizing to avoid installation and functionality issues.

This document is organized around small, medium, and large events-per-second workloads. For example, each day, your environment might have thousands of events per second. However, the total workload depends not only on the event data received through SmartConnectors or ArcSight Enterprise Security Manager (ESM), but also other workloads that occur at the same time. For example someone might be searching for events, new information can be coming in about the entities associated with the events, system backup operations are performed, and a various other operations might occur. The system must be able to process all of these types of transactions simultaneously and maintain satisfactory performance.

The conditions in your environment are likely to be somewhat different than in our test lab and, as such, it is possible you might need to further adjust the system sizing or tuning values for satisfactory performance in your environment.

Database Cluster

Your deployment can have one of the following database clusters:

- **Collocated database cluster** - In a collocated database cluster, the database, master node, and worker node are deployed on a single node.

  The collocated database cluster has been used to determine the system sizing and tuning for On-Premises single-node deployment - small and medium workloads.
- **Non-collocated database cluster** - In a non-collocated database cluster, the database is not deployed on the worker nodes in the CDF cluster. Instead, the database is deployed on dedicated nodes that make up the database cluster, and this cluster is not a part of the CDF cluster.

  The non-collocated database cluster has been used to determine the system sizing and tuning for the following scenarios:
  - On-premises multi-node deployment - large workload
  - Cloud multi-node deployment - small, medium, and large workloads.

For both these clusters, the communal storage is a separate entity and we have dedicated sizing for the communal storage and the database node(s). The following are the communal storages for each of the deployment scenarios:

- On-Premises deployment - MinIO
- Azure deployment - Azure Blob
- AWS deployment - S3

Small Workload

As we complete testing for additional scenarios, we will add more information to this page.

This section describes the system sizing and tuning for the ArcSight Platform and deployed capabilities Transformation Hub, Fusion, Command Center for ESM, Intelligence, Recon, and the ArcSight Database that has been confirmed in our testing lab to maintain satisfactory performance of the system under a small workload.

Workloads

This section describes the workload that was placed on the tested system.

Event Workload

Application	Events per second
Microsoft Windows	700
InfoBlox NIOS	700
Intelligence Data (VPN, AD, Proxy)	100
Total	1,500

Other workload

Category	Level
Storage Groups	10
Searches	3 per hour (concurrent)
Reports	1 scheduled every hour

System Sizing

This section describes the system sizing of the tested system.

On-Premises Deployment

The "CDF Master/Worker Node/Database Node" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, Recon, and Database compute components were deployed in an all-in-one collocated configuration on the tested system. However, the Database "Communal Storage" components were deployed on a separate node because they are not embedded within the ArcSight Platform. When using this information as guidance for your own system sizing, the "CDF Master/Worker Node/Database Node" system resources are always needed, but the Database "Communal Storage" system resources are only needed when deploying Recon or Intelligence.

Category	1 x CDF Master/Worker Node/Database Node	1 x Communal Storage
Processor	Intel(R) Xeon(R) Gold 6248 CPU @ 2.50GHz	Intel(R) Xeon(R) Gold 6248 CPU @ 2.50GHz
vCPU(s) (# threads)	24	6
RAM (per node)	128 GB	32 GB
Disks (per node)	ESX data store	ESX data store
Storage per day (1x)	7 GB (depot) + 15 GB (ES)	27 GB (MinIO)
Total disk space (5 Billion events)	1 TB (holds upto 45 days of events)	1 TB (holds upto 40 days of events)
K-safety level	0	N/A

AWS Deployment

The "CDF Worker (Platform)" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, and Recon components were deployed on the tested system. However, Intelligence components were deployed on the "CDF Worker (Intelligence)" system resources because they utilize a significant amount of resources when running analytics jobs. When using this information as guidance for your own system sizing, the "CDF Worker (Platform)" system resources are always needed, the "Database" system resources are only needed when deploying Recon or Intelligence, and the "CDF Worker (Intelligence)" system resources are only needed when deploying ArcSight Intelligence.

Category	CDF Worker (Platform)	Database	CDF Worker (Intelligence)
Instance Type	m5.2xlarge	m5d.4xlarge	m5.2xlarge
Instance Count	3	3	3
Disks (per node)	500 GB - EBS storage (gp2)	2 x 300 NVMe SSD	500 GB - EBS storage (gp2)

Azure Deployment

The "CDF Worker (Platform)" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, and Recon components were deployed on the tested system. However, Intelligence components were deployed on the "CDF Worker (Intelligence)" system resources because they utilize a significant amount of resources when running analytics jobs. When using this information as guidance for your own system sizing, the "CDF Worker (Platform)" system resources are always needed, the "Database" system resources are only needed when deploying Recon or Intelligence, and the "CDF Worker (Intelligence)" system resources are only needed when deploying ArcSight Intelligence.

Category	CDF Worker (Platform)	Database	CDF Worker (Intelligence)
Instance Type	D2s_V3	D4s_V3	D2s_V3
Instance Count	3	3	3
Disks (per node)	1 x 500 GB - Premium SSD	2 x 300 GB - Premium SSD	1 x 500 GB - Premium SSD

System Tuning

This section describes the system tuning of the tested system.

- Database Tuning
- Transformation Hub Tuning
- Intelligence Tuning
- Fusion Tuning
- SmartConnector Tuning

Database Tuning

Category	Property	On-Premises	AWS	Azure
Core Database	shard_count	3	3	3
Core Database	depot_size	40%	60%	60%

Tuple Mover	tm_concurrency	5	6	5
Tuple Mover	tm_memory	10G	10G	10G
Tuple Mover	plannedconcurrency	5	6	5
Tuple Mover	tm_memory_usage	10000	10000	10000
Tuple Mover	maxconcurrency	10	7	10
Ingest Resource pools	ingest_pool_memory_size	30%	30%	30%
Ingest Resource pools	ingest_pool_planned_concurrency	6	6	6
Backup	Backup Interval (hours)	1	1	1
Communal Storage	Server-side Encryption	Disabled	Disabled	Yes (MMK)

Transformation Hub Tuning

Property	On-Premises	AWS	Azure
# of Kafka broker nodes in the Kafka cluster	1	3	3
# of ZooKeeper nodes in the ZooKeeper cluster	1	3	3
# of Partitions assigned to each Kafka Topic*	12	24	24
# of replicas assigned to each Kafka Topic	1	2	2
# of message replicas for the __consumer_offsets Topic	1	3	3
Schema Registry nodes in the cluster	1	3	3
# of CEF-to-Avro Stream Processor instances to start**	0	0	0
# of Enrichment Stream Processor Group instances to start	2	2	2

*Kafka topics - th-arcsight-avro, mf-event-avro-enriched, and, if connectors are configured to send to Transformation Hub in CEF format, th-cef

**If connectors are configured to send Avro format to Transformation Hub, you can set the quantity to 0 because there is no need to convert CEF to Avro.

Intelligence Tuning

Property	On-Premises	AWS	Azure
Elasticsearch Shard Count	6	6	6
Elasticsearch data processing Instances	1	3	3
Elasticsearch Index Replica Count	0	1	1

Elasticsearch Memory (GB)	10	4	4
Elasticsearch number of cores	6	2	2
Elasticsearch Size Per Batch	5mb	5mb	5mb
Logstash Instances	2	3	3
Logstash pipeline workers per instance	2	1	1
Logstash Pipeline Batch size	500	500	500
LogStash Filter Applied	yes	yes	yes
Spark parallelism	32	32	32
Spark number of executors	3	3	3
Spark executor memory	5g	4g	4g
Spark number of executor cores	1	1	1
Spark driver memory	4g	4g	3g
Spark memory overhead factor	0.2	0.2	0.2
Intelligence Job per day	1	1	1

Fusion Tuning

Category	All Deployments
Event Integrity Check Task Count	1
Event Integrity Check Chunk Size	1000
Use Event Integrity Check Resource Pool	false

SmartConnector Tuning

Category	All Deployments
SmartConnector Version	8.3.0.14008.0
Instance Count	1
Acknowledgement Mode	leader
usessl (Transformation Hub Destination Param)	false
contenttype (Transformation Hub Destination Param)	Avro
topic (Transformation Hub Destination Param)	th-arcsight-avro
compression.type	gzip
transport.batchqueuesize	20000

transport.cefkafka.batch.size	50000
transport.cefkafka.linger.ms	10
transport.cefkafka.max.request.size	4194304
transport.cefkafka.multiplekafkaproducers	true
transport.cefkafka.threads	3

Medium Workload

As we complete testing for additional scenarios, we will add more information to this page.

This section describes the system sizing and tuning for the ArcSight Platform and deployed capabilities Transformation Hub, Fusion, Command Center for ESM, Intelligence, Recon, and the ArcSight Database that has been confirmed in our testing lab to maintain satisfactory performance of the system under a medium workload.

Workloads

This section describes the workload that was placed on the tested system.

Event Workload

This table provides event ingestion workload in events per second:

Application	On-Premises Collocated Database	Azure Non-Collocated Database
Microsoft Windows	2,400	9,000
InfoBlox NIOS	2,400	9,000
Intelligence Data (VPN, AD, Proxy)	250	2,000
Total	5,050	20,000

Other Workload

Category	Level
Storage Groups	10
Searches	3 per hour (concurrent)
Reports	1 scheduled every hour

System Sizing

This section describes the system sizing of the tested system.

- On-Premises Deployment
- Azure Deployment
- AWS Deployment

On-Premises Deployment

The "CDF Master/Worker Node/Database Node" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, Recon, and Database compute components were deployed in an all-in-one collocated configuration on the tested system. However, the Database "Communal Storage" components were deployed on a separate node because they are not embedded within the ArcSight Platform. When using this information as guidance for your own system sizing, the "CDF Master/Worker Node/Database Node" system resources are always needed, but the Database "Communal Storage" system resources are only needed when deploying Recon or Intelligence.

Category	1 x CDF Master/Worker Node/Database Node	1 x Communal Storage
Processor	Intel(R) Xeon(R) Gold 6248 CPU @ 2.50GHz	Intel(R) Xeon(R) Gold 6248 CPU @ 2.50GHz
vCPU(s) (# threads)	48	8
RAM (per node)	192 GB	48 GB
Disks (per node)	ESX data store	ESX data store
Storage per day (1x)	10 GB (depot) + 20 GB (ES)	100 GB (MinIO)
Total disk space (5 Billion events)	1 TB (holds up to 30 days of events)	1 TB (holds up to 15 days of events)
K-safety level	0	N/A

Azure Deployment

The "CDF Worker (Platform)" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, and Recon components were deployed on the tested system. However, Intelligence components were deployed on the "CDF Worker (Intelligence)" system resources because they utilize a significant amount of resources when running analytics jobs. When using this information as guidance for your own system sizing, the "CDF Worker (Platform)" system resources are always needed, the "Database" system resources are only

needed when deploying Recon or Intelligence, and the "CDF Worker (Intelligence)" system resources are only needed when deploying ArcSight Intelligence.

Category	CDF Worker (Platform)	Database	CDF Worker (Intelligence)
Instance Type	D16s_V3	D32s_V3	D16s_V3
Instance Count	3	6	3
Disks (per node)	2 TB - Premium SSD	2 TB (Depot) - Premium SSD	2 TB - Premium SSD

AWS Deployment

The "CDF Worker (Platform)" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, and Recon components were deployed on the tested system. However, Intelligence components were deployed on the "CDF Worker (Intelligence)" system resources because they utilize a significant amount of resources when running analytics jobs. When using this information as guidance for your own system sizing, the "CDF Worker (Platform)" system resources are always needed, the "Database" system resources are only needed when deploying Recon or Intelligence, and the "CDF Worker (Intelligence)" system resources are only needed when deploying ArcSight Intelligence.

Category	CDF Worker (Platform)	Database	CDF Worker (Intelligence)
Instance Type	m5.4xlarge	m5.12xlarge	m5.4xlarge
Instance Count	3	6	3
Disks (per node)	1 X 2048 GB (gp3) EBS Volumes	8 x 250 GB (gp3) EBS Volumes	1 X 2048 GB (gp3) EBS Volumes

System Tuning

This section describes the system tuning of the tested system.

- Database Tuning
- Transformation Hub Tuning
- Intelligence Tuning
- Fusion Tuning
- SmartConnector Tuning

Database Tuning

Category	Property	On-Premises	Azure	AWS
Core Database	shard_count	3	18	18

Core Database	depot_size	40%	60%	60%
Tuple Mover	tm_concurrency	5	5	10
Tuple Mover	tm_memory	10G	10G	10G
Tuple Mover	plannedconcurrency	5	5	5
Tuple Mover	tm_memory_usage	10000	10000	20000
Tuple Mover	maxconcurrency	10	10	10
Ingest Resource pools	ingest_pool_memory_size	30%	30%	30%
Ingest Resource pools	ingest_pool_planned_concurrency	6	6	6
Backup	Backup Interval (hours)	1	1	1

Transformation Hub Tuning

Property	On-Premises	Azure	AWS
# of Kafka broker nodes in the Kafka cluster	1	3	3
# of ZooKeeper nodes in the ZooKeeper cluster	1	3	3
# of Partitions assigned to each Kafka Topic*	12	72	72
# of replicas assigned to each Kafka Topic	1	2	2
# of message replicas for the __consumer_offsets Topic	1	3	3
Schema Registry nodes in the cluster	1	3	3
# of CEF-to-Avro Stream Processor instances to start**	0	0	3
# of Enrichment Stream Processor Group instances to start	2	3	3

*Kafka topics - th-arcsight-avro, mf-event-avro-enriched, and, if connectors are configured to send to Transformation Hub in CEF format, th-cef

**If connectors are configured to send Avro format to Transformation Hub, you can set the quantity to 0 because there is no need to convert CEF to Avro.

Intelligence Tuning

Property	On-Premises	Azure	AWS
Elasticsearch Shard Count	6	6	6
Elasticsearch data processing Instances	1	3	3
Elasticsearch Index Replica Count	0	1	1
Elasticsearch Memory (GB)	14	12	12
Elasticsearch number of cores	8	6	5

Elasticsearch Size Per Batch	5mb	5mb	5mb
Logstash Instances	3	12	15
Logstash pipeline workers per instance	2	2	1
Logstash Pipeline Batch size	500	1000	500
LogStash Filter Applied	yes	yes	yes
Spark Parallelism	32	32	64
Spark number of executors	3	8	9
Spark executor memory	6g	8g	7g
Spark number of executor cores	1	1	1
Spark Driver Memory	6g	8g	8g
Spark Memory Overhead Factor	0.2	0.2	0.2
Intelligence Job per day	1	1	1

Fusion Tuning

Category	All Deployments
Event Integrity Check Task Count	1
Event Integrity Check Chunk Size	1000
Use Event Integrity Check Resource Pool	false

SmartConnector Tuning

Category	All Deployments
SmartConnector Version	8.3.0.14008.0
Instance Count	1
Acknowledgement Mode	none
usessl (Transformation Hub Destination Param)	false
contenttype (Transformation Hub Destination Param)	Avro
topic (Transformation Hub Destination Param)	th-arcsight-avro
compression.type	gzip
transport.batchqueuesize	20000
transport.cefkafka.batch.size	50000
transport.cefkafka.linger.ms	10

transport.cefkafka.max.request.size	4194304
transport.cefkafka.multiplekafkaproducers	true
transport.cefkafka.threads	6
syslog.parser.threadcount	6

Large Workload

As we complete testing for additional scenarios, we will add more information to this page.

This section describes the system sizing and tuning for the ArcSight Platform and deployed capabilities Transformation Hub, Fusion, Command Center for ESM, Intelligence, Recon, and the ArcSight Database that has been confirmed in our testing lab to maintain satisfactory performance of the system under a large workload.

Workloads

This section describes the workload that was placed on the tested system.

Event Workload

Application	Events per second
Microsoft Windows	54,000
InfoBlox NIOS	54,000
Intelligence Data (VPN, AD, Proxy)	12,000
Total	120,000

Other Workload

Category	Level
Storage Groups	10
Searches	3 per hour (concurrent)
Reports	1 scheduled every hour

System Sizing

This section describes the system sizing of the tested system.

AWS Deloyment

The "CDF Worker (Platform)" system resources are where the core platform, Transformation Hub, Fusion, Command Center for ESM, and Recon components were deployed on the tested system. However, Intelligence components were deployed on the "CDF Worker (Intelligence)" system resources because they utilize a significant amount of resources when running analytics jobs. When using this information as guidance for your own system sizing, the "CDF Worker (Platform)" system resources are always needed, the "Database" system resources are only needed when deploying Recon or Intelligence, and the "CDF Worker (Intelligence)" system resources are only needed when deploying ArcSight Intelligence.

Category	CDF Worker (Platform)	Database	CDF Worker (Intelligence)
Instance Type	m5.8xlarge	m5d.8xlarge	m5.8xlarge
Instance Count	3	18	6
Disks (per node)	3 TB - EBS storage	2 x 600 NVMe SSD	3 TB - EBS storage

System Tuning

This section describes the system tuning of the tested system.

- Database Tuning
- Transformation Hub Tuning
- Intelligence Tuning
- Fusion Tuning
- SmartConnector Tuning

Database Tuning:

Category	Property	AWS
Core Database	shard_count	18
Core Database	depot_size	60%
Tuple Mover	tm_concurrency	5
Tuple Mover	tm_memory	10G
Tuple Mover	plannedconcurrency	5
Tuple Mover	tm_memory_usage	10000
Tuple Mover	maxconcurrency	10

Ingest Resource pools	ingest_pool_memory_size	30%
Ingest Resource pools	ingest_pool_planned_concurrency	6
Backup	Backup Interval (hours)	1
Communal Storage	Server-side Encryption	disabled

Transformation Hub Tuning

Property	AWS
# of Kafka broker nodes in the Kafka cluster	3
# of ZooKeeper nodes in the ZooKeeper cluster	3
# of Partitions assigned to each Kafka Topic*	108
# of replicas assigned to each Kafka Topic	2
# of message replicas for the __consumer_offsets Topic	3
Schema Registry nodes in the cluster	3
# of CEF-to-Avro Stream Processor instances to start**	0
# of Enrichment Stream Processor Group instances to start	6

*Kafka topics - th-arcsight-avro, mf-event-avro-enriched, and, if connectors are configured to send to Transformation Hub in CEF format, th-cef

**If connectors are configured to send Avro format to Transformation Hub, you can set the quantity to 0 because there is no need to convert CEF to Avro.

Kafka Override Parameters	AWS
arcsight.eventbroker.kafka.KAFKA_NUM_IO_THREADS	256
arcsight.eventbroker.kafka.KAFKA_NUM_NETWORK_THREADS	52
arcsight.eventbroker.kafka.KAFKA_NUM_REPLICA_FETCHERS	145

Intelligence Tuning

Property	AWS
Elasticsearch Shard Count	6
Elasticsearch data processing Instances	6
Elasticsearch Index Replica Count	1
Elasticsearch Memory (GB)	24
Elasticsearch number of cores	12

Elasticsearch Size Per Batch	10mb
Logstash Instances	108
Logstash pipeline workers per instance	2
Logstash Pipeline Batch size	2000
Spark Parallelism	64
Spark number of executors	24
Spark executor memory	12g
Spark number of executor cores	1
Spark Driver Memory	8g
Spark Memory Overhead Factor	0.2
Intelligence Job per day	1

Fusion Tuning

Category	All Deployments
Event Integrity Check Task Count	6
Event Integrity Check Chunk Size	1000
Use Event Integrity Check Resource Pool	false

SmartConnector Tuning

Category	All Deployments
SmartConnector Version	8.3.0.14008.0
Instance Count	5
Acknowledgement Mode	none
usessl (Transformation Hub Destination Param)	false
contenttype (Transformation Hub Destination Param)	Avro
topic (Transformation Hub Destination Param)	th-arcsight-avro
compression.type	gzip
transport.batchqueuesize	20000
transport.cefkafka.batch.size	50000
transport.cefkafka.linger.ms	10

transport.cefkafka.max.request.size	4194304
transport.cefkafka.multiplekafkaproducers	true
transport.cefkafka.threads	6

File System Options

This section describes the available file system options.

- Network File System Options
- Network File System Minimum Directory Sizes

Network File System Options

The following table lists the minimum network file system (NFS) options

Category	Minimum Requirement
NFS Types	• Amazon EFS • HPE 3PAR File Persona • Linux-based NFS • NetApp
NFS Server Versions	• NFSv4.1 You might enable additional versions on the NFS server, however ArcSight Platform only uses NFSv4.1.

Network File System Minimum Directory Sizes

The following table lists the minimum required size for each of the NFS installation directories.

Directory	Minimum Size	Description
{NFS_ROOT_DIRECTORY}/itom-vol	130 GB	This is the CDF NFS `root` folder, which contains the CDF database and files. The disk usage will grow gradually.
{NFS_ROOT_DIRECTORY}/db-single-vol	Start with 10 GB	This volume is only available when you did not choose PostgreSQL High Availability (HA) for CDF database setting. It is for CDF database. During the install you will not choose the Postgres database HA option.

{NFS_ROOT_DIRECTORY}/db-backup-vol	Start with 10 GB	This volume is used for backup and restore of the CDF Postgres database. Its sizing is dependent on the implementation's processing requirements and data volumes.
{NFS_ROOT_DIRECTORY}/itom-logging-vol	Start with 40 GB	This volume stores the log output files of CDF components. The required size depends on how long the log will be kept.
{NFS_ROOT_DIRECTORY}/arcsight-volume	10 GB	This volume stores the component installation packages.

Install File System Options

The following table lists the supported file system options.

Category	Requirement
ext4	—
xfs	ftype=1

Firewall Ports

When you install ArcSight Platform and deploy the associated capabilities, you will need to open firewall ports for many of the elements that make up the Platform. For more information about the ports for each component, see "Understanding Firewall Ports for the ArcSight Platform" in the *Administrator's Guide to ArcSight Platform 22.1*.

Examples of Deployment Scenarios

You can deploy the ArcSight Platform capabilities in a variety of ways. The most basic deployment option is an all-in-one system that contains a limited number of capabilities on a single node. The single-node deployment is suitable for small workloads or to use as a proof-of-concept environment. For large workloads, you will need a multi-node environment, possibly with multiple masters. There are many scenarios and considerations involved in creating your environment. Please see "Prerequisites and Considerations for Adding Capabilities" in the *Administrator's Guide for ArcSight Platform*.

This section provides some examples on how you could deploy one or more capabilities. Use these examples as a general guidance for planning your environment.

Multiple Master and Worker Nodes for High Availability

In this scenario, which **deploys Intelligence with high availability**, you have three master nodes connected to three worker nodes and a database cluster. Each node runs on a separate, dedicated, connected host. All nodes have the same operating system, such as CentOS 7.8. Each Worker Node processes events, with failover to another Worker Node if a Worker fails. All of these environments require an external server to support NFS.

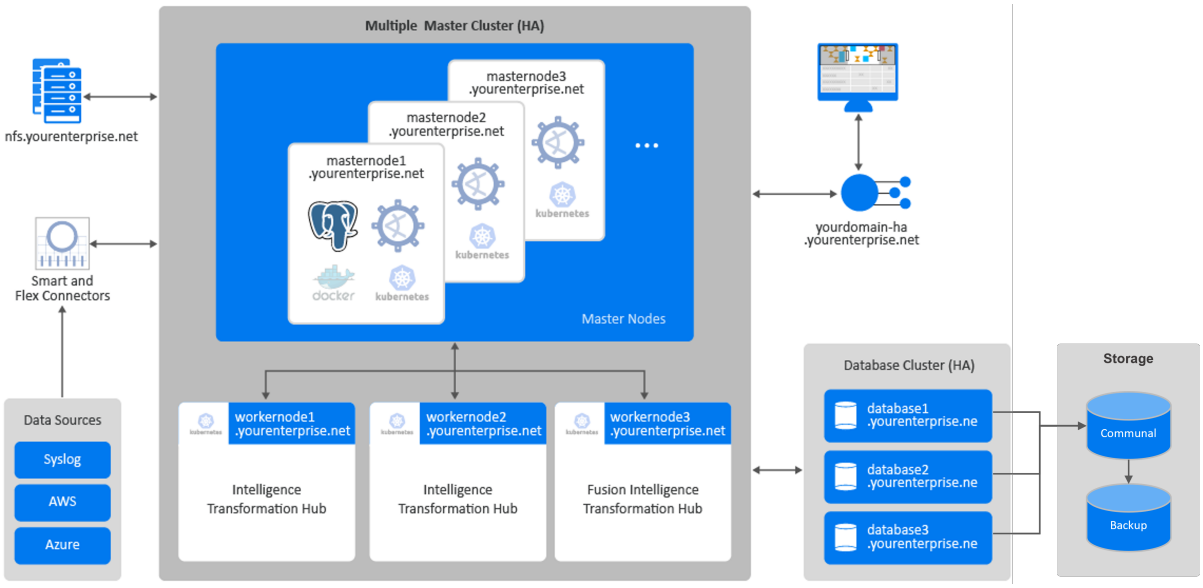
You can run this configuration in development and testing. It is the recommended configuration for highly available environments.

> ✔ If this scenario resembles your intended deployment, you might want to use the `example-install-config-intelligence-high_availability.yaml` config file with the ArcSight Platform Installer. See "Configuring the Deployed Capabilities" in the *Administrator's Guide for ArcSight Platform*.

The worker nodes process events, with failover to another worker node in the event of a worker failure. There are no single points of failure. You need a minimum of nine physical or VM environments: three dedicated master nodes, three or more dedicated worker nodes, and a database cluster. You also need a customer-provisioned, highly available NFS server (external NFS).

Diagram of this Scenario

Figure 1. *Example deployment of Intelligence in a high-availability cluster*



Characteristics of this Scenario

This scenario has the following characteristics:

- The Kubernetes cluster has three master nodes and three worker nodes, so that it can tolerate a failure of a single master and still maintain master node quorum.
- A FQDN hostname for a virtual IP is used so that clients accessing master nodes have a single reliable hostname to connect to that will shift to whatever is the current primary master node. For example, `yourdomain-ha.yourenterprise.net`.
- Transformation Hub's Kafka and ZooKeeper are deployed to all worker nodes with data replication enabled (1 original, 1 copy) so that they can tolerate a failure of a single node and still remain operational.
- Intelligence services, as well as Transformation Hub's platform and processing services, are allocated across all worker nodes so that, if one of the nodes fails, Kubernetes can move all of the components to the other node and still remain operational.
- Fusion is allocated to a single worker node.
- For the NFS configuration, use an NFS server that has high availability capabilities so that it is not a single point of failure.
- The database cluster has three nodes with data replication enabled (1 original and 1 copy) so that it can tolerate a failure of a single node and remain operational.

Guidance for Node Configuration

The following table provides guidance for deploying the capabilities across multiple nodes to support a large workload.

Node Name	Description	RAM	CPU Cores	Disk Space	Ports
Master Nodes 1-3 `masternodeNN.yourenterprise.net`	CDF Management Portal	256 GB	32	5 TB	CDF Vault CDF Management Portal Kubernetes NFS
Database Nodes 1-3 `databaseNN.yourenterprise.net`	Database	192 GB	24	28 TB	Database
Worker 1 `workernode1.yourenterprise.net`	Intelligence Transformation Hub	256 GB	32	5 TB	Kubernetes Transformation Hub
Worker 2 `workernode1.yourenterprise.net`	Intelligence Transformation Hub	256 GB	32	5 TB	Kubernetes Transformation Hub
Worker 3 `workernode1.yourenterprise.net`	Fusion Intelligence Transformation Hub	256 GB	32	5 TB	ArcMC Intelligence Kubernetes Transformation Hub

Single Master, Multiple Workers, and a High-availability Database

In this scenario, which **deploys Intelligence with high availability on the ArcSight Database**, you have a single master node connected to three worker nodes and a database cluster. This scenario supports an environment with modest EPS and minimal number of nodes. However, it allows for futher scaling with multiple worker nodes. Each worker node runs on a separate, dedicated, connected host. All nodes have the same operating system, such as CentOS 7.8.

- **Diagram of this Scenario**
- **Characteristics of this Scenario**
- **Guidance for Node Configuration**

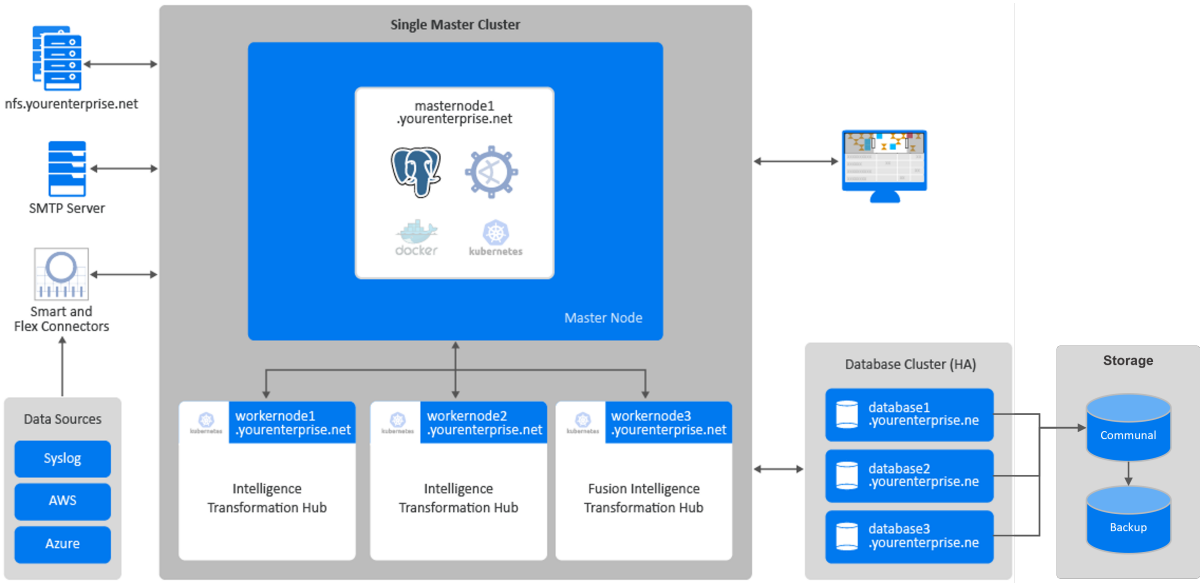
You can run this configuration in development and testing. This the recommended configuration for having a highly available database.

If this scenario resembles your intended deployment, you might want to use the `example-install-config-intelligence-scale_db.yaml` config file with the ArcSight Platform Installer. See "Configuring the Deployed Capabilities" in the *Administrator's Guide for ArcSight Platform*.

You need a minimum of nine physical or VM environments: three dedicated master nodes, three or more dedicated worker nodes, and a database cluster. You also need a customer-provisioned, highly-available NFS server (External NFS) and an SMTP server.

Diagram of this Scenario

Figure 1. *Example deployment of Intelligence and Recon*



Charcteristics of this Scenario

This scenario has the following characteristics:

- The Kubernetes cluster overall is not highly available since it is deployed with only one master node.
- A FQDN hostname for a virtual IP is used so that clients accessing master nodes have a single reliable hostname to connect to that will shift to whatever is the current primary master node. For example, `yourdomain-ha.yourenterprise.net`.
- Transformation Hub's Kafka and ZooKeeper are deployed to all worker nodes with data replication enabled (1 original, 1 copy) so that they can tolerate a failure of a single node and still remain operational.
- Transformation Hub's ZooKeeper is deployed to all worker nodes with data replication across the nodes so that it can tolerate a failure of a single node and still remain operational.
- Intelligence services, Fusion, and Transformation Hub's platform and processing services are allocated across all worker nodes so that, if one of the nodes fails, Kubernetes can move all of the components to the other node and still remain operational.
- The database cluster has three nodes with data replication enabled (1 original and 1 copy) so that it can tolerate a failure of a single node and remain operational.

- For the NFS configuration, use an NFS server that has high availability capabilities so that it is not a single point of failure.

Guidance for Node Configuration

The following table provides guidance for deploying the Intelligence across multiple nodes to support a medium workload.

Node Name	Description	RAM	CPU Cores	Disk Space	Ports
Master Node `masternode1.yourenterprise.net`	CDF Management Portal (Optional) Fusion	256 GB	32	5 TB	CDF Management Portal Kubernetes NFS
Database Nodes 1-3 `databaseNN.yourenterprise.net`	Database	192 GB	24	28 TB	Database
Worker 1 `workernode1.yourenterprise.net`	Intelligence Fusion Transformation Hub	256 GB	32	5 TB	ArcMC Intelligence Kubernetes Transformation Hub
Worker 2 `workernode2.yourenterprise.net`	Intelligence Fusion Transformation Hub	256 GB	32	5 TB	ArcMC Intelligence Kubernetes Transformation Hub
Worker 3 `workernode3.yourenterprise.net`	Fusion Intelligence Transformation Hub	256 GB	32	5 TB	ArcMC Intelligence Kubernetes Transformation Hub

Everything on a Single Node

In this scenario, which **deploys ESM Command Center on a single node**, you have the master and worker node co-located. You can include ArcSight SOAR as an optional capability on the same node.

- Diagram of this Scenario
- Scenario Characteristics
- Guidance for Node Configuration

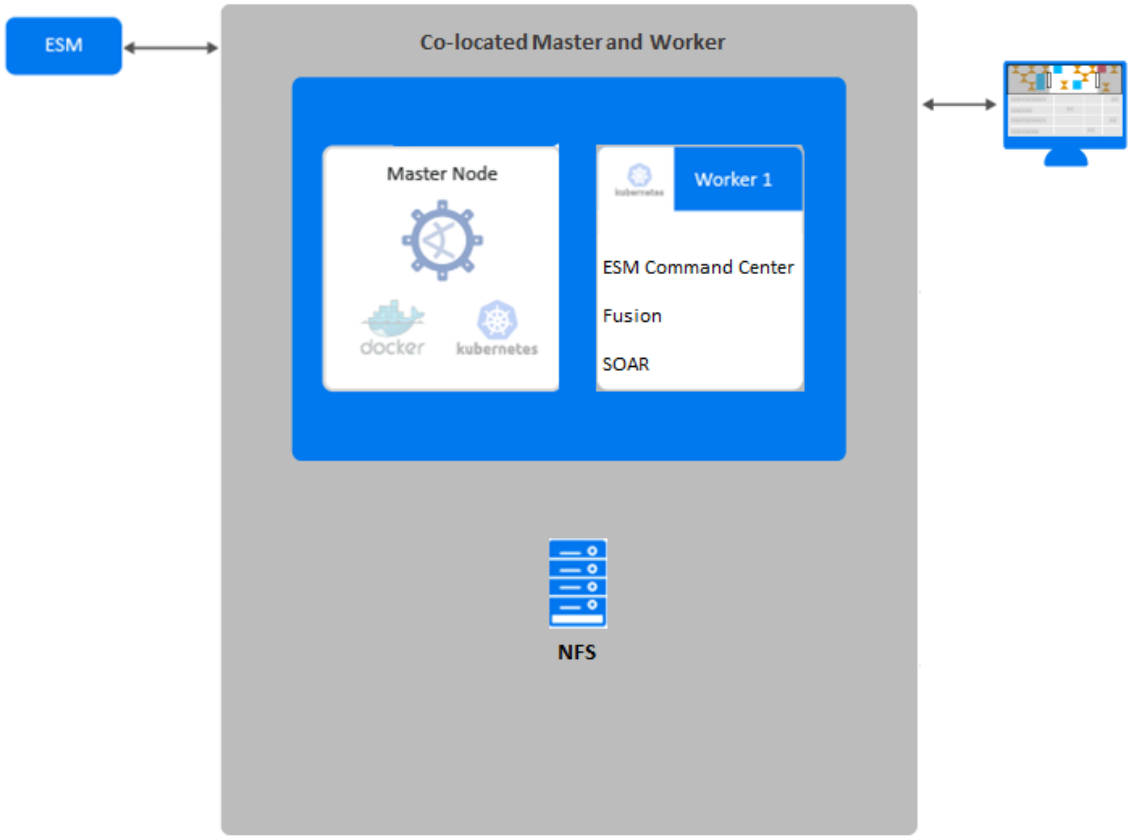
You can run this configuration in development and testing environments.

If this scenario resembles your intended deployment, you might want to use the `example-install-config-esm_cmd_center-single-node.yaml` config file with the ArcSight Platform Installer. See "Configuring the Deployed Capabilities" in the *Administrator's Guide for ArcSight Platform*. The configuration in the example file describes a single-node deployment, but you can add more worker nodes to the file.

You need a minimum of one physical or VM environment to support master, worker, and NFS server on a single node. If you intend to install ESM Manager on the same machine, install ESM Manager first. ESM Manager uses port 8443, so master-api-ssl-port is set to a different port to avoid a conflict.

Diagram of this Scenario

Figure 2. *Example deployment of ESM Command Center on a single node*



Characteristics of this Scenario

This scenario has the following characteristics:

- The Kubernetes cluster has a single node to which you deploy ESM Command Center, Fusion, and (optionally) SOAR.

  Having a single master node creates a single point of failure. As a result, if you intend to add worker nodes, this configuration is not recommended for high availability (HA) environments.

- FIPS 140 mode is enabled.
- For the NFS configuration, an NFS server that has high availability capabilities so that it is not a single point of failure.

Guidance for Node Configuration

The following table provides guidance for deploying ESM Command Center and associatd capabilities on a single node to support a small workload.

Node Name	Description	RAM	CPU Cores	Disk Space	Ports
Master Node `yourdomain-node.yourenterprise.net`	CDF Management Portal	256 GB	32	5 TB	CDF Vault CDF Management Portal Kubernetes NFS
Worker 1	ESM Command Center Fusion SOAR (optional)	32 GB	8	300 GB	ArcMC Kubernetes SOAR

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 ArcSight Platform Technical Requirements (ArcSight Platform 22.1)

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 Documentation-Feedback@microfocus.com.

We appreciate your feedback!

Legal Notices

Copyright Notice