Cisco Secure Network Analytics Data Store Design Guide
# Table of Contents

**Audience** ........................................................................................................................... 3  
**Purpose of this Document** .................................................................................................. 3  
**Executive Summary** .......................................................................................................... 3  
**Technology Overview** ........................................................................................................ 4  
Manager .................................................................................................................................. 4  
Flow Collector .......................................................................................................................... 4  
UDP Director ........................................................................................................................... 5  
Data Store ............................................................................................................................... 5  
**Architecture Overview** ....................................................................................................... 6  
Traditional Distributed Architecture ..................................................................................... 6  
Central Data Store Architecture ............................................................................................ 6  
Data Resiliency ........................................................................................................................ 7  
Network Connectivity ............................................................................................................. 8  
Hardware Data Store .............................................................................................................. 8  
Virtual Data Store .................................................................................................................. 8  
Flow Ingestion .......................................................................................................................... 8  
Traffic Profiles ........................................................................................................................ 9  
Sizing Virtual Data Store Node Deployments ......................................................................... 10  
Data Store Sizing Calculation: ............................................................................................... 10  
**Requirements and Considerations** .................................................................................. 11  
Licensing Requirements .......................................................................................................... 11  
Hardware Networking Requirements .................................................................................... 12  
Virtual Networking Requirements ......................................................................................... 12  
Caveats as of version 7.3.1 ..................................................................................................... 13  
**Designs and Recommendations** ....................................................................................... 14  
Hardware or Virtual ................................................................................................................ 14  
Standard Design - Hardware ................................................................................................. 15  
Management Communication ................................................................................................. 15  
Inter-node Communication ..................................................................................................... 15  
Power Redundancy .................................................................................................................. 17  
Out of Band Management ...................................................................................................... 17  
Adding Redundancy and Scale to the Hardware Design ......................................................... 18  
Redundancy ............................................................................................................................. 18  
Multiple ESX/KVM Host Design - Virtual ............................................................................ 20  
Single Host ESX/KVM Design - Virtual ................................................................................ 22  
Adding Redundancy and Scale to Virtual Designs ................................................................. 23
Redundancy .............................................................................................................................. 23
Retention and Scalability ........................................................................................................ 24
Software Installation and Configuration ................................................................................. 25
    Initial Setup Tool .................................................................................................................. 25
    Appliance Setup Tool .......................................................................................................... 25
    SystemConfig ..................................................................................................................... 25
Manager Data Store Configuration .......................................................................................... 26
Summary .................................................................................................................................. 27
Reference .................................................................................................................................. 28
    Communication Ports ........................................................................................................... 28
Design Check-List .................................................................................................................... 29
Audience

The audience for this document includes sales engineers, field consultants, professional services specialists, IT and security managers, partner engineers, and customers who want to take advantage of an infrastructure built to deliver efficiency and enable innovation.

Purpose of this Document

This document provides the fundamental design principles needed to understand how the Secure Network Analytics (formerly Stealthwatch) Data Store scales telemetry consumption, provides data resiliency and increases search performance for the most demanding enterprise class environments.

Executive Summary

The Cisco Secure Network Analytics Data Store solution provides a central repository to store network telemetry collected by Flow Collectors. The Data Store is comprised of a cluster of Data Nodes, each containing a portion of the network’s flow data, and a backup copy of a separate Data Node’s data. Because all of the flow data is in one centralized database, as opposed to spread across multiple Flow Collectors, the Secure Network Analytics Manager can retrieve query results from a single Data Store faster than querying all Flow Collectors individually. The Data Store database cluster gives customers improved fault tolerance, improved query response, faster graph and chart population when using Cisco Secure Network Analytics. The Data Store architecture allows for independent scaling of flow collection and flow storage functions to better fit customer needs and budgets.
Technology Overview

Cisco Secure Network Analytics provides enterprise-wide network visibility and applies advanced security analytics to detect and respond to threats in real time. Using a combination of behavioral modeling, machine learning, and global threat intelligence, Secure Network Analytics can quickly, and with high confidence, detect threats such as command-and-control (C&C) attacks, ransomware, distributed denial-of-service (DDoS) attacks, illicit cryptomining, unknown malware, and insider threats. With a single, agentless solution, Cisco Secure Network Analytics delivers comprehensive threat monitoring across the entire network, for both clear and encrypted communications.

The Secure Network Analytics Data Store was first introduced in release 7.3.0 as a hardware appliance offering. In the following 7.3.1 release the Secure Network Analytics Data Store was expended to also include virtual appliances on VMWare’s ESXi hosts and KVM hosts. A Secure Network Analytics Data Store deployment can be either a physical appliance deployment consisting of a Manager (SMC2210), at least one Data Store (DS6200) and minimally one Flow Collector (FC4210) or a virtual appliance deployment of a Manager VE, three Data Node VE’s and minimally one Flow Collector VE. Mixing physical hardware appliances with virtual appliances is not supported in the 7.3.1 release but is road mapped for the near future.

Manager

The Secure Network Analytics Manager aggregates, organizes, and presents analysis from Flow Collectors, the Identity Services Engine, and other sources. It uses graphical representations of network traffic, identity information, customized summary reports, and integrated security and network intelligence to deliver comprehensive security analytics.

The Manager also serves as the Central Manager for all Secure Network Analytics appliances. The Manager is the centralized user interface in which all configuration changes, management and investigative workflows take place.

<table>
<thead>
<tr>
<th>PID</th>
<th>Description</th>
<th>Form Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-SMC2210-K9</td>
<td>Stealthwatch Management Console 2200</td>
<td>1RU</td>
</tr>
<tr>
<td>L-ST-SMC-VE-K9</td>
<td>Stealthwatch Management Console VE</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Flow Collector

The Flow Collector leverages enterprise telemetry such as NetFlow, IPFIX (Internet Protocol Flow Information Export), and other types of flow data from existing infrastructure such as routers, switches, firewalls, endpoints, and other network infrastructure devices. The Flow Collector can receive enhanced flow from ETA enabled exporters as well as collect telemetry from proxy data sources. This enhanced flow and proxy data can be analyzed by the cloud-based, multilayered machine learning engine to deliver zero-day threat detection.

<table>
<thead>
<tr>
<th>PID</th>
<th>Description</th>
<th>Form Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-FC4210-K9</td>
<td>Stealthwatch Flow Collector 4210</td>
<td>1RU</td>
</tr>
<tr>
<td>L-ST-FC-VE-K9</td>
<td>Stealthwatch Flow Collector VE</td>
<td>N/A</td>
</tr>
</tbody>
</table>
UDP Director

The UDP Director simplifies the collection and distribution of network and security data across the enterprise. It helps reduce the processing power on network routers and switches by receiving essential network and security information from multiple locations and then forwarding it to a single data stream to one or more destinations.

<table>
<thead>
<tr>
<th>PID</th>
<th>Description</th>
<th>Form Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-UDP2210-K9</td>
<td>Stealthwatch UDP Director 2210</td>
<td>1RU</td>
</tr>
<tr>
<td>L-ST-UDP-VE-K9</td>
<td>Stealthwatch UDP Director VE</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Data Store

The Secure Network Analytics Data Store provides a central repository to store a network’s telemetry, collected by Flow Collectors. The Data Store is comprised of a three or more Data Nodes, with each Data Node containing a portion of ingested data, and a backup of a separate Data Node’s data. Instead of data being spread across multiple flow collectors and databases, all network telemetry is consolidated in the centralized database. This centralized architecture is optimized to provide fast query results, rapid graph and chart population and long-term data storage. Additionally, the Data Store improves fault tolerance, offering the highest level of data redundancy and enterprise class deployment resiliency.

This new database architecture is a major improvement to the existing distributed model. The following sections will explore the technical details of this evolutionary architecture to explain how it operates, and the various deployment designs so that you can determine the one that works best for your specific environment.

<table>
<thead>
<tr>
<th>PID</th>
<th>Description</th>
<th>Form Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-DS6200-K9</td>
<td>Stealthwatch Data Store 6200</td>
<td>6RU</td>
</tr>
<tr>
<td>ST-DS6200-D1-K9</td>
<td>Stealthwatch Data Store Spare Node</td>
<td>2RU</td>
</tr>
<tr>
<td>L-ST-DS-VE-K9</td>
<td>Data Node VE</td>
<td>N/A</td>
</tr>
</tbody>
</table>
**Architecture Overview**

**Traditional Distributed Architecture**

In a traditional Secure Network Analytics deployment, one or more Flow Collectors ingests network telemetry, deduplicates flow data, performs analysis, and reports directly to the Manager. To resolve user-submitted queries, including graphs and charts, the Manager queries all of the managed Flow Collectors. Each flow collector returns matching results to the Manager. The Manager collates the information from the different result sets, then generates a graph or chart displaying the results. In this deployment, each Flow Collector stores flow data on a local database. See the following diagram for an example.

![Diagram of Traditional Distributed Architecture](image)

This Distributed Architecture poses design challenges when deployments scale across many Flow Collectors or when flow rates are multiple millions of flows per second (FPS).

**Central Data Store Architecture**

In a Secure Network Analytics deployment with a central Data Store architecture, the Data Store database cluster sits logically between your Manager and Flow Collectors. A Secure Network Analytics Data Store is comprised of at least one DS6200 consisting of three physical Data Nodes, or three Data Node VE’s. A physical Data Store deployment can support up to 12 DS6200’s (maximum of 36 Data Nodes). If the deployment of multiple DS6200’s is required to meet flow ingest or retention requirements, each DS6200 appliance set will combine to form a single Data Store. In version 7.3.1 adding more than three Data Node VE’s to a virtual Data Store deployment is not supported.

In a central Data Store architecture, one or more Flow Collectors ingests, deduplicates, and performs analysis on flow data. The processed telemetry is then sent directly to the Data Store for storage. The Data Store collates the information from the different data sets, storing it within the Data Store database cluster, and distributing it equally to all of the Data Nodes within the Data Store database cluster.

![Diagram of Central Data Store Architecture](image)
This centralized design allows the Manager to send all queries to a single destination instead of querying each Flow Collector individually. In addition to overall greater retention times, the Data Store architecture enables faster overall query performance. The Data Store supports an optional cold spare Data Node, which can be added to the Data Store’s database in the event of a failure with one of the managed Data Nodes.

**Data Resiliency**

One of the improvements coming with the Secure Network Analytics Data Store and the new centralized architecture is the implementation of K-safety functionality.

K-safety enables the fault tolerance in the Secure Network Analytics Data Store database cluster. The value K represents the number of times the data in the database cluster is replicated. These replicas allow other nodes to take over query processing for any failed nodes, which results in no data loss upon an unexpected node failure, allowing processing and ingestion to continue during a failure event.

Each single node will write a copy of its own tables onto its associated buddy node to ensure resiliency across the database.

A Secure Network Analytics Data Store deployment with a single DS6200 (i.e., three Data Nodes) or three single Data Node VE’s will have a K-Safety of 1, meaning it is possible to lose up to 1 Data Node out of the 3 total Data Nodes without interruption to query actions or the storage of data. This is because each Data Node’s tables are written to a single peer to create a mirror copy Data Node’s primary table. When a Secure Network Analytics Data Store is expanded by adding more nodes the K-Safety feature will allow for up to 50% of nodes failing while maintaining complete functionality.

See examples of each design and the associated Data Node redundancy below:

### Single DS6200

![Single DS6200 Diagram]

The green square represents the primary database table, while the grey square represents the replicated copy in case of Data Node failure.

In a Secure Network Analytics Data Store design using a cold-spare Data Node, the Data Store will be setup and configured as a typical deployment. The DS6200’s 3 Data Nodes will be added to the Data Store and the database will be initialized without the inclusion of the spare Data Node. The cold-spare Data Node (DS6200-D1) will be managed by Central Manager just as the other nodes, but it will not be added to the Data Store’s database until a Data Node has failed.

At the time of a failure the spare Data Node will be manually added to the Data Store’s database. Once added the Data Store will replicate the data from the existing Data Nodes restoring the databases to normal operation.

### Two DS6200s

![Two DS6200s Diagram]

### Single DS6200 + Cold Spare

![Single DS6200 + Cold Spare Diagram]
Network Connectivity

Hardware Data Store

For hardware deployments, there are two interface configurations for handling appliance Management Communications and Flow Ingestion. The configuration will be made during the initial setup on the Manager (SMC2210), Flow Collector (FC4210), and each Data Node (DS6200).

Physical Network Connectivity:

- Secure Network Analytics Data Store appliances support either a Base-T (copper) or SFP (fiber) configuration for Management Communication and Flow Ingestion.
- The Manager and Flow Collector can be configured with a 10-Gbps Ethernet SFP+ (default) or 1-Gb/10-Gb BASE-T Ethernet LAN port (RJ45 connector).
- Each Data Node can be configured with a 10-Gbps Ethernet SFP+ (default) or 10-Gb BASE-T Ethernet LAN port (RJ45 connector).

While each appliance can be configured independently, it is recommended to configure all appliances using the same media. It is also recommended to deploy the appliances using 10Gbps connections.

In deployment scenarios where Flow Collectors are not co-located with the Data Store, it will be necessary to ensure adequate bandwidth exists between them to support Data Store updates from the Flow Collector. Optimum design will allow for a minimum of 4Gbps of bi-directional burst traffic between these appliances.

Virtual Data Store

It is recommended that all network adapter types are set to VMXNET 3 to obtain optimal performance. If Secure Network Analytic virtual appliances are distributed across multiple ESXi/KVM hosts it is recommended that vnic uplinks are 10Gbps, redundant, and NIC teaming is implemented.

Data Node devices will need two vnic s one for management communication and the other for internode communication.

Flow Ingestion

Getting flow out of a network and into a Flow Collector engine is an early step in establishing a successful Secure Network Analytics deployment. Typically, for most Secure Network Analytic deployments, NetFlow will be enabled at the access, distribution, and core to give administrators the entire conversation both north-south and east-west.

Each network is different and as such special consideration is needed when designing how to get NetFlow into the Flow Collector.

There are several types of telemetry exporters can provide that are supported by Secure Network Analytics. Supported telemetry protocols include:

- NetFlow version 5 – Non-customizable, not template based format
- NetFlow version 9 – Customizable, template-based format
- IPFIX – Optimal Telemetry format for NetOps events and SecOps detections
- Sampled NetFlow – Sampled NetFlow is true NetFlow sent at a sampled rate, 1 out of 1000 packets for example
- sFlow – Supported on some non-Cisco devices, useful for NetOps, less so for SecOps
- SYSLOG – Standard message for logging, can add context to flow records such as URL data
- NSEL – NetFlow Security Event Logging is provided by the Cisco ASA and FTD, the protocol includes fields such as pre and post network address translations.
Typically, NetFlow is collected from the WAN edge, campus core and access layers from a wide range of exporters: switches, routers, firewalls, wireless LAN controllers and flow sensors.

Most exporters only support two destinations for exporting telemetry and if there are multiple tools with collectors needing this data there is a potential for a constraint on where telemetry can be sent. This can be solved by deploying a UDP Director. With the UDP Director all telemetry can be exported to a single destination and configured with granular forwarding rules to ensure each collector is receiving the data it requires.

Traffic Profiles

Telemetry profiles should be taken into consideration when setting the retention period for the Data Store. Factors to consider include the flow per second rate, types of traffic profiles, and requirements for how long it is necessary to store the data.

**Standard Enterprise traffic profile:**
This is the typical environment that closely resembles a campus architecture with standard NetFlow enabled exporters. In these environments the Flow Collectors typically consolidate flows from five or more exporters as the network communication traverses the campus network. This traffic profile is most common and suitable for most designs.

### Physical Appliance Sizing Table for Standard Traffic Profile

<table>
<thead>
<tr>
<th>Flow Rate (FPS)</th>
<th># FC 4210</th>
<th># DS6200 for 90 days storage</th>
<th># DS6200 for 180 days storage</th>
<th># DS6200 for 360 days storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>250K</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>500K</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>1M</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>2M</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>3M</td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>-</td>
</tr>
</tbody>
</table>
Service Provider traffic profile:
Service Providers have unique environments consisting of high numbers of unique hosts, very little deduplication of flows and large amounts of sampled flow data. Due to this profile more appliances are required for flow collection and data storage to meet the flow rate and retention times. The Secure Network Analytics Data Store supported up to 90 days of storage when using the service provider traffic profile.

Physical Appliance Sizing Table for Service Provider Traffic Profile

<table>
<thead>
<tr>
<th>Flow Rate (FPS)</th>
<th># FC 4210</th>
<th># DS6200 for 30 days storage</th>
<th># DS6200 for 60 days storage</th>
<th># DS6200 for 90 days storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>250k</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>500K</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>1M</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>2M</td>
<td>8</td>
<td>4</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>3M</td>
<td>12</td>
<td>6</td>
<td>12</td>
<td>-</td>
</tr>
</tbody>
</table>

Sizing Virtual Data Store Node Deployments
While physical appliances have fixed resources, and are easily scaled and sized using the previous tables, virtual Data Store deployments may require additional calculations to support specific retention times.

In version 7.3.1 expanding the size of the Data Store post initial configuration is not supported. It is recommended to plan for appropriate Data Store growth when sizing a virtual Data Store deployment.

Data Store Sizing Calculation:

\[(\text{Avg FPS} / 1000) \times 1.6 \times \text{Retention days} / \text{DN#} = \text{Data Storage required per Data Node}\]

1. Divide the FPS by 1000
2. Multiply this number by 1.6 GB. Note: A day’s worth of storage at an ingress rate of 1,000 FPS is 1.6GB
3. Multiply this number by the required retention time in number of days
4. Divide this number by 3, which is the number of Data Nodes supported in a virtual Data Store

For example, if the environment being monitored produces 50,000 daily average FPS and the customer needs storage for 90 days with 3 Data Nodes the calculation will work as following:

\[(50,000/1000) \times 1.6 \times 90 / 3 = 2400 \text{ GB (2.4 TB) per Data Node}\]

1. Daily average FPS = 50,000
2. 50,000 average FPS / 1,000 = 50
3. 50 x 1.6 GB = 80 GB for one day’s worth of flow storage
4. 80 GB x 90 days = 7200 GB of total storage for 90 days
5. 7200 GB of total storage / 3 Data Nodes = 2400 GB (2.4 TB) per Data Node

30 days 220 x 16 = 2200 + 1320 = 3,520 GB
90 days 220 x 16 = 2200 + 1320 = 3520 x3 = 10,560 GB
# Requirements and Considerations

## Licensing Requirements

Your Secure Network Analytics deployment requires a Flow Rate (FPS) Smart License; the Data Store itself does not require an additional license.

## Basic Hardware Specifications

<table>
<thead>
<tr>
<th>Appliance</th>
<th>Platform</th>
<th>Gen</th>
<th>Height</th>
<th>SW Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS6200</td>
<td>UCS-C240 (3x)</td>
<td>M5SX</td>
<td>6 RU (3x 2 RU)</td>
<td>7.3.0 and above</td>
</tr>
<tr>
<td>SMC2210</td>
<td>UCS-C220</td>
<td>M5SX</td>
<td>1 RU</td>
<td>7.3.0 and above</td>
</tr>
<tr>
<td>FC4210</td>
<td>UCS-C220</td>
<td>M5SX</td>
<td>1 RU</td>
<td>7.3.0 and above</td>
</tr>
</tbody>
</table>

When configuring a Secure Network Analytics Data Store deployment all Secure Network Analytics appliances must be on the same version.

## Manager VE

<table>
<thead>
<tr>
<th>Users</th>
<th>Memory Reserved</th>
<th>CPU Reserved</th>
<th>Storage</th>
<th>SW Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 9</td>
<td>32 GB</td>
<td>4</td>
<td>125 GB</td>
<td>7.3.1 and above</td>
</tr>
<tr>
<td>10 or more</td>
<td>64 GB</td>
<td>8</td>
<td>200 GB</td>
<td>7.3.1 and above</td>
</tr>
</tbody>
</table>

## Flow Collector VE

<table>
<thead>
<tr>
<th>Flows Per Second</th>
<th>Memory Reserved</th>
<th>CPU Reserved</th>
<th>Storage</th>
<th>SW Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 50,000</td>
<td>32 GB</td>
<td>6</td>
<td>200 GB</td>
<td>7.3.1 and above</td>
</tr>
<tr>
<td>Up to 120,000</td>
<td>70 GB</td>
<td>8</td>
<td>200 GB</td>
<td>7.3.1 and above</td>
</tr>
</tbody>
</table>

## Data Store VE

<table>
<thead>
<tr>
<th>Flows Per Second</th>
<th>Memory Reserved</th>
<th>CPU Reserved</th>
<th>Storage</th>
<th>SW Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 50,000</td>
<td>32 per Data Node</td>
<td>6 per Data Node</td>
<td>800 GB per Data Node 2.4 TB total Data Store</td>
<td>7.3.1 and above</td>
</tr>
<tr>
<td>Up to 120,000</td>
<td>32 GB per Data Node</td>
<td>12 per Data Node</td>
<td>1.92 TB per Data Node 5.76 TB Data Store</td>
<td>7.3.1 and above</td>
</tr>
<tr>
<td>Up to 220,000</td>
<td>64 GB per Data Node</td>
<td>16 per Data Node</td>
<td>3.52 TB per Data Node 10.56 TB Data Store</td>
<td>7.3.1 and above</td>
</tr>
</tbody>
</table>

© 2021 Cisco and/or its affiliates. All rights reserved.
Hardware Networking Requirements
Take consideration of the following requirements for management and inter-node communications when designing Data Store deployment using hardware appliances:

Management communication requirements:
1. 10Gbps throughput is required for the Data Store’s management interface, which can be BASE-T or SFP.
2. Clock skew must be kept to 1 second or lower between and among your Data Nodes.
3. When using 10G throughput ensure there is 4 Gbps of bandwidth available between each FC4210 and the Data Store’s Data Nodes.

Inter-node communication requirements:
1. 10 Gbps full duplex switched layer-2 connection exists between the Data Node and supports 6.4Gbps of throughput or greater.
2. Round-trip time (RTT) latency between Data Nodes must not exceed 200 microseconds.
3. Clock skew must not exceed 1 second between Data Nodes.

Virtual Networking Requirements
Take consideration of the following requirements for management and inter-node communications when designing Data Store deployment using virtual appliances:

Management communication requirements:
In a design where Secure Network Analytic devices are deployed across ESX/KVM hosts it is required that the physical adapter uplink that is used for management communications is 10G.

In this example there is a defined port group named Management Port Group. The vnic eth0 interfaces for each Secure Network Analytic Device are assigned to the Management Port Group. It is required that the interfaces in the management port group and its associated vswitch has at least one 10G uplink. It is a best practice to ensure that more than one 10G uplink is available for redundancy and NIC teaming is implemented.

Inter-node communication requirements:
For a virtual deployments with data nodes spread across multiple ESX/KVM hosts it is required that the physical adapter uplink for each inter-node eth1 vnic is 10Gbps.

It is a best practice to ensure that more than one 10G uplink is available for redundancy and NIC teaming is implemented. Ensure the inter-node eth1 vnic is assigned a distributed or standard port group tagged with a VLAN that has been configured on the networking device uplink. This ensures each internode interface on each Data Node can communicate with the corresponding inter-node interfaces on the other Data Nodes within their private network.
Caveats as of version 7.3.1

Take into consideration the following limitations and features that are not supported for the 7.3.1 release of the Secure Network Analytics Data Store:

- Java Client Support / Desktop Client
- Support for Endpoint License
- Fed Certification (FIPS, DODINAPL, CC, Gv6)
- Support for Virtual Appliances is limited to a total of three Data Nodes
- The DS6200 hardware appliances require the FC4210 hardware appliance for telemetry ingest
- Mixing hardware and virtual appliances is not supported
- Applications
  - Visibility Assessment
  - Host Classifier
  - Cryptographic Audit
Designs and Recommendations

Hardware or Virtual

The first decision that to make in any Data Store design is whether or not to deploy the solution physically or virtually. There are considerations to be made that each customer opportunity will define. Perhaps the customer is resistant to deploying new hardware or maybe the customer needs a hardware deployment in order to handle their FPS ingest rates.

The largest factor for many customers will be the scale of the deployment.

If the design calls for ingest rates above 220k FPS then the solution should be designed for hardware. For networks that produce less than 220k FPS it is likely the virtual solution will make the most sense for the customer.

Placing the Secure Network Analytics Data Store

As a repository for flow data collected by the Flow Collector, and as the centralized repository against which a Secure Network Analytics Manager runs queries, install the Data Nodes at a location in the network that is accessible by the Flow Collector and the Secure Network Analytics Manager.

It is common in Enterprise networks to deploy Secure Network Analytics on a management network. In this design exporters will need network connectivity to the management interface on the Flow Collector. Telemetry from these network exporters will then be de-duplicated and processed by the Flow Collector’s engine and stored as database entries in the Data Store.

There is an additional design consideration to allow the Flow Collector to collect telemetry from standalone or isolated networks by using a second collection interface on the Flow Collector. This can be beneficial in a design where a test lab or a DMZ is segmented away from the production network.

This design makes it possible to physically segment all management communications from other isolated areas of the network.
Standard Design - Hardware

Let's start with the simplest design where a standard single DS6200 is deployed with one SMC, Flow Collector, and single Nexus switch. After which, we'll consider other configurations and design choices which provide increased redundancy and scale.

This design will include the following Secure Network Analytics appliances:

- SMC2210
- FC4210
- DS6200

Management Communication

For best performance the Secure Network Analytics Manager should be co-located with the Data Nodes to ensure the optimal performance of queries. Each Secure Network Analytics appliance must have a routable IP address assigned to the eth0 management port.

Secure Network Analytics appliances all have a logical interface eth0 that is used for encrypted management and data communications between all appliances. This is also how configuration changes are pushed by the Manager to other appliances as well as queries made to the Data Store by the Manager.

When deploying physical Data Store appliances, the logical eth0 management port can be configured to use a BASE-T copper 1G/10G port or SFP+/DAC 10G port. While each appliance is configured independently, it is recommended to use the same media type and 10Gbps connectivity between all Secure Network Analytics appliances.

The Secure Network Analytics Manager establishes management of all the other Secure Network Analytics appliances in the ecosystem using Central Manager.

It is recommended that all Secure Network Analytics appliances including each Data Node in the Data Store deployment be on the same network (i.e., Layer-3 subnet).

It is possible to have Secure Network Analytics appliances on different networks as long as they can reach the management (i.e., eth0) interface of the other appliances.

Ensure all ports are open that are necessary for inter-device communications. See the Communication Ports section at the end of the document for more information.

The Secure Network Analytics appliances must be managed by the same Manager. Ensure each appliance uses the same domain name server and NTP (Network Time Protocol) source.

Inter-node Communication

The Data Nodes are required to be co-located with each other to minimize latency for critical inter-node communication. Each CDS Node will communicate on a private non-routable LAN that will be used for the inter-node communication. Traffic sent between Data Nodes is not encrypted.

Cisco recommends the use of the Nexus 7000 and Nexus 9000 series switches for the inter-node communication switch.
This design features a single link using the local eth2 interface on each Data Node to communicate with each Data Nodes in the deployment.

Each Data Node will use this link for writing data and other critical database communications.

There are two management configuration options with this design, the default which uses the SFP port only, and an alternative which uses the two onboard RJ-45 ports and the first two SFP ports.

By default the 10-Gbps Ethernet SFP+ ports are used for all network communication. In this configuration the left most SFP port is the eth0 management port and the third SFP+ port from the left is used for inter-node communication to a single inter-node communication switch as indicated below:

![Diagram showing eth0 and eth2 interfaces](image1)

The other option is to use the 1-Gb/10-Gb BASE-T Ethernet LAN port (RJ45 connector) for management and the left most 10-Gbps Ethernet SFP+ for inter-node communication to a single inter-node communication switch as illustrated:

![Diagram showing eth2 and eth0 interfaces](image2)

For additional details on appliance specifications refer the Data Store spec sheet:
Power Redundancy
Each CDS Node will be shipped with redundant PSUs labelled in the diagram below. This will ensure you can maintain power redundancy in case of an outage.

<table>
<thead>
<tr>
<th>PID</th>
<th>Description</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-DS6200</td>
<td>Stealthwatch Data Store 6200</td>
<td>Redundant [770W or 1050 W] AC 50/60 Auto Ranging (100V to 240V)</td>
</tr>
</tbody>
</table>

Out of Band Management
Out of band management can be configured using the Cisco Integrated Management Controller (CIMC) port. This port is labeled with a symbol on the back of the appliance as indicated in the picture below.

For further details on configuring CIMC for remote access please follow the guidance found in the Cisco UCS Installation Guide:

Adding Redundancy and Scale to the Hardware Design

Redundancy

Dual-switch architecture - In this redundant switch configuration, the inter-node communications have redundancy via a dual-link port channel using both the eth2 and eth3 ports on each Data Node.

This configuration is completed in the Initial Setup Tool on each Data Node in the Data Store.

Redundant SMC 2210 - Adding a redundant Manager is the most common form for redundancy in Secure Network Analytics. The architecture includes a total of two Managers with one serving as the primary and the other as the secondary. The secondary operates in the pair as read only.

Configuration changes made to the primary are automatically synched to the secondary. If the primary Manager goes down, the secondary must be manually promoted to primary before it is fully operational.
Redundant FCNF 4210 – To prevent losing the ability to process flow data in the event a Flow Collector goes down, you can add multiple FCNF 4210’s for redundancy. The flow data would be sent to both Flow Collectors and both Flow Collectors would be managed by the same Manager. A flow that is sent to both collectors will count against the flow rate license twice. Keep this in mind when building a design.

Additionally, a UDP Director would be recommended for the redundant flow collector option to prevent having to configure multiple destinations from the exporters.

High-Availability UDP Director 2210 – NetFlow is sent to a virtual IP address shared between two UDP Directors in the high-availability design for the UDP Director. One UDP Director will be the primary online device and the other UDP director will be the secondary offline device. If the primary node should fail, the secondary node takes over automatically and becomes the primary.

Two physical 1Gbps copper cross connect cables are required between the appliances. The device must be located within a range of each no more than 300 feet.

Retention and Scalability

Expanding the size of the Data Store – Expanding the size of the Data Store is accomplished by adding additional DS6200. Recall the DS6200 is comprised of three Data Nodes. The Secure Network Analytics Data Store supports a maximum of 12 DS6200s, which equates to 36 Data Nodes in total. Adding additional DS6200s to the cluster requires the Data Nodes to be registered with the Central Manager and additional database setup steps to bring them online.

For detailed steps on how to expand the Data Store find the documentation linked below:

Scaling Flow Collection – Bottlenecks can occur when flow rates exceed a collector’s maximum flows per second thresholds causing the Flow Collector to drop telemetry packets. To prevent loss of telemetry data a UDP Director 2210 and additional FC4210’s can be deployed. The recommended approach is to plan a maintenance window where the Secure Network Analytics Flow Collector can be taken offline. During this window the UDP Director is configured to use the IP address of the existing single Flow Collector, preventing unnecessary changes to production switches and routers. The original Flow Collector and new Flow Collectors can be uniquely IP addressed and then flow forwarding rules can be defined to distribute the flows per second load across multiple collectors.
Multiple ESX/KVM Host Design – Virtual

Overview
For virtual Data Store designs using multiple ESX/KVM hosts, it is recommended to deploy each virtual Data Node on a separate ESX/KVM host to ensure data resiliency across the Data Store. It also can help spread required virtual resources such as compute and memory across multiple hosts. Below is a recommended design where each virtual Data Node is distributed across three different ESX/KVM hosts.

In this design it is recommended to deploy a primary and secondary Manager and two Flow Collectors. Having a primary and secondary Manager ensures that if an ESX/KVM host was to fail there would always be one Manager online. Having two flow collectors is important if a customer wants to store flow redundantly or needs multiple collectors to scale their flows per second ingest.

This design is the most resilient and provides for higher flow rates but requires the most resources. Take this into consideration for your design.

Management Communication
Ensure that the mgmt eth0 vnic on each Secure Network Device is in a port group tagged with the same VLAN across each ESX/KVM host. It is recommended to use a distributed vSwitch and a distributed port group if possible.
Ensure that the network device each uplink is connected to and its associated switchport interface has the correct VLAN configured to ensure all devices can communicate over their management interface.

Inter-node Interface Configuration

Ensure that the inter-node eth1 vnic on each Data Node is in a port group tagged with the same VLAN across each ESX/KVM host. It is recommended this port group is created on a separate vswitch that is used only by the inter-node vnics.

Additionally it is recommended to have redundant 10G uplinks. Ensure that the network device each uplink(s) is physically connected to and its associated switchport has the correct VLAN configured. This will ensure internode vnics can communicate over their network.
Single Host ESX/KVM Design – Virtual

Overview

For lower FPS designs or if there are environmental resource constraints the Secure Network Analytic devices can be deployed on a single ESX/KVM host. For these circumstances the following design is recommended.

This is the easiest design to deploy and configure, but with all Secure Network Analytic devices stored on one ESX/KVM host, you will not have data resiliency if the ESX/KVM host fails.

In the event a single Data Node VE virtual machine was to fail, data resiliency would be maintained.

Management Communication

Ensure that the mgmt eth0 vnic on each Secure Network Device is minimally on the same vswitch. It is recommended to segment these vnics into a port group with a dedicated vswitch.

This will ensure other VM’s and their activity such as snapshots or intensive network activity does not impact the Secure Network Analytic devices.

Inter-node Interface Configuration

Ensure that the inter-node eth1 vnic on each Data Node is in a port group tagged with the same VLAN across each ESX/KVM host. It is recommended this port group is created on a separate vswitch that is used only by the inter-node vnics.

It is recommended to have redundant 10G uplinks. Ensure that the network device each uplink(s) is physically connected to and its associated switchport has the correct VLAN configured. This will ensure internode vnics can communicate over their network.
Adding Redundancy and Scale to Virtual Designs

Redundancy

**Redundant Manager VE** - Adding a redundant Manager is the most common form for redundancy in Secure Network Analytics. The architecture includes a total of two Managers with one serving as the primary and the other as the secondary. The secondary operates in the pair as read only.

Configuration changes made to the primary are automatically synched to the secondary. If the primary Manager goes down, the secondary must be manually promoted to primary before it is fully operational.

**Normal Operation Scenario**

- SMC Primary
- SMC Secondary
- Management and DB Entries
- Data Store VE
- Flow Collector VE

**Fault Recovery Scenario**

- SMC Primary
- SMC Secondary
- Management and DB Entries
- Data Store VE
- Flow Collector VE

In virtual deployments deployed across multiple ESX/KVM hosts it is recommended to deploy a Manager on two different hosts. This ensures in the event of a failure to one of the ESX/KVM hosts an SMC will be available.

**Redundant Flow Collector VE** - To prevent losing the ability to process flow data in the event a Flow Collector goes down, you can add multiple Flow Collector VE’s for redundancy. The flow data would be sent to both Flow Collectors and both Flow Collectors would be managed by the same Manager. A flow that is sent to both collectors will count against the flow rate license twice. Keep this in mind when building a design.

Additionally, a UDP Director would be recommended for the redundant flow collector option to prevent having to configure multiple destinations from the exporters.
Retention and Scalability

**Scaling Flow Collection** – Bottlenecks can occur when flow rates exceed a collector’s maximum flows per second thresholds causing the Flow Collector to drop telemetry packets. To prevent loss of telemetry UDPVE’s and additional Flow Collector VE’s can be deployed. The recommended approach is to plan a maintenance window where the Secure Network Analytics Flow Collector can be taken offline. During this window the UDP Director is configured to use the IP address of the existing single Flow Collector, preventing unnecessary changes to production switches and routers. The original Flow Collector and new Flow Collectors can be uniquely IP addressed and then flow forwarding rules can be defined to distribute the flows per second load across multiple collectors.
Software Installation and Configuration

Initial Setup Tool

The Initial Setup Tool will be used to configure the interfaces used and the management interface IP address settings. For simplicity it is recommended to use the same subnet when assigning management IP addresses for each appliance.

For the Manager and Flow Collector appliances there is an additional step to choose whether or not the device will be part of a Data Store deployment. Once the configuration has been completed you must RFD the appliance to change the Data Store selection.

For the Data Node appliances, the IST will have an additional option to configure the interface configuration and IP address settings for the inter-node communications. There are two interface configuration options for hardware Data Nodes. The first being a single link and the second being a port channel.

Once the interface configuration is chosen the IP addressing will be configured. IP addressing must reside within the 169.254.42.0/24 subnet.

Note: If the root or sysadmin passwords are forgotten, the Recovery option within SystemConfig can be used to reset the passwords.

Appliance Setup Tool

After the Initial Setup Tool the next step is to complete the Appliance Setup Tool (AST). The AST is accessed via a web browser by navigating to the management IP address of each appliance. The appliances must be configured in priority order starting with the Manager. Following it, the AST must be completed on each Data Node then the Flow Collector(s).

The Appliance Setup Tool is responsible for configuring the following areas:

1. Verifying the password settings and changing the admin web interface password
2. Validating the management interface IP address settings
3. Defining the device’s host name and domain
4. Configuring the domain name server(s) the device will use
5. Providing the NTP source(s) the device will use
6. Register the device with Central Manager, this will be the SMC2210.

After completing the AST each device will reboot.
SystemConfig

If the root or sysadmin passwords are forgotten, the Recovery option within SystemConfig can be used to reset the passwords.

Manager Data Store Configuration

Following the IST and AST there is a final necessary step to configure the communication between the Manager and the Data Nodes. The configuration is completed on the Manager from the SystemConfig GUI and consists of completing the Datastore Secure Connectivity setup tool.

The outcome of this step is a secure, SSL certificate based, communication channel between the Manager and the Data Nodes is established and database user accounts will be configured.

For more information on the IST, AST, and Manager Data Store Configuration:
Summary

Effective network security analytics is not a function of applying just one technique. To stay ahead of evolving threats, a network visibility and analytics solution needs to be able to use a combination of methods. This begins by collecting and storing telemetry data at scale for comprehensive visibility on which analytical techniques such as behavioral modeling and machine learning can be applied.

Today with the new Secure Network Analytics Data Store customers can now meet the needs of the world's most demanding and dynamic enterprise-class and service provider networks. The Data Store architecture allows for ingesting up to 3M FPS and telemetry from over 30M endpoints which can be redundantly stored for more than half a year.

This new architecture transforms the capabilities of existing customers and offers a great solution to new customers who demand the fastest query response times and require data stored at scale for extended periods of time.

Learn more

Product and solution pages: Cisco SecureX Network Analytics Cisco SecureX Cloud Analytics Encrypted Traffic Analytics Free visibility assessment
Cisco highly recommends placing the Data Nodes within your firewall. Below is a list of the ports you will need to have open to ensure connectivity between the Secure Network Analytics Appliances.

<table>
<thead>
<tr>
<th>From (Client)</th>
<th>To (Server)</th>
<th>Port</th>
<th>Protocol</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manager</td>
<td>Flow Collectors and Data Nodes</td>
<td>22/TCP</td>
<td>SSH, TCP</td>
<td>SSH, required to initialize Data Store database</td>
</tr>
<tr>
<td>Manager</td>
<td>Flow Collectors and Data Nodes</td>
<td>443/TCP</td>
<td>HTTPS, TCP</td>
<td>HTTPS, required for secure communications between appliances</td>
</tr>
<tr>
<td>Manager</td>
<td>Data Nodes</td>
<td>5444/TCP</td>
<td>HTTPS, TCP</td>
<td>Vertica Management Console secure communications</td>
</tr>
<tr>
<td>Flow Collectors</td>
<td>Manager</td>
<td>443/TCP</td>
<td>HTTPS, TCP</td>
<td>HTTPS, required for secure communications between appliances</td>
</tr>
<tr>
<td>sFlow Exporters</td>
<td>Flow Collectors – sFlow</td>
<td>6343/UDP</td>
<td>UDP</td>
<td>sFlow ingestion</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>All other Data Nodes</td>
<td>22/TCP</td>
<td>SSH, TCP</td>
<td>SSH, required to initialize Data Store database and for database administration tasks</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>Manager</td>
<td>443/TCP</td>
<td>HTTPS, TCP</td>
<td>HTTPS, required for secure communications between appliances</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>All other Data Nodes</td>
<td>4803/TCP</td>
<td>TCP</td>
<td>Inter-Data Node messaging service</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>All other Data Nodes</td>
<td>4804/UDP</td>
<td>UDP</td>
<td>Inter-Data Node messaging service</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>All other Data Nodes</td>
<td>5433/UDP</td>
<td>UDP</td>
<td>Vertica Messaging service monitoring</td>
</tr>
<tr>
<td>Data Nodes</td>
<td>All other Data Nodes</td>
<td>6543/UDP</td>
<td>UDP</td>
<td>Inter-Data Node messaging service</td>
</tr>
<tr>
<td>All Appliances</td>
<td>Data Nodes</td>
<td>5433/TCP</td>
<td>TCP</td>
<td>Vertica client connections</td>
</tr>
<tr>
<td>NTP Service</td>
<td>Manager, Flow Collectors, and Data Nodes</td>
<td>123/UDP</td>
<td>UDP</td>
<td>NTP, required for time synchronization</td>
</tr>
<tr>
<td>Administrator workstations</td>
<td>Manager, FC and Data Nodes</td>
<td>443/TCP</td>
<td>TCP</td>
<td>WebUI Management</td>
</tr>
<tr>
<td>Administrator workstations</td>
<td>Manager, FC and Data Nodes</td>
<td>22/TCP</td>
<td>TCP</td>
<td>Management/Troubleshooting</td>
</tr>
</tbody>
</table>
Design Check-List

Virtual or Hardware?

Which interface will be used for management on the Manager?
10-Gbps Ethernet SFP+ (default)
1-Gb/10-Gb BASE-T Ethernet LAN port (RJ45 connector)

Which interface will be used for management on each Flow Collector?
10-Gbps Ethernet SFP+ (default)
1-Gb/10-Gb BASE-T Ethernet LAN port (RJ45 connector)

Which interface will be used for management on each Data Node?
10-Gbps Ethernet SFP+ (default)
1-Gb/10-Gb BASE-T Ethernet LAN port (RJ45 connector)

Management IP Addressing:
Manager: (Ex: 192.168.1.100/24)
DS Node1: (Ex: 192.168.1.111/24)
DS Node2: (Ex: 192.168.1.112/24)
DS Node3: (Ex: 192.168.1.113/24)
Flow Collector: (Ex: 192.168.1.200/24)

Inter-node communication IP Addressing:
(MUST be within the 169.254.42.0/24 subnet)
DS Node1: (Ex: 169.254.42.11/24)
DS Node2: (Ex: 169.254.42.12/24)
DS Node3: (Ex: 169.254.42.13/24)

Passwords:
(CL) root:
(CL) sysadmin:
(web) admin:

Domain Name Servers:
DNS Server1:
DNS Server2:

Network Time Protocol Servers:
NTP Server1:
NTP Server2:
NTP Server3:

Certificate Configurations:
If CA specific certs are required for device identity certificates, they need to be generated for each device and added via Central Manager.