HITACHI
Inspire the Next

Hitachi NAS Platform

NAS Platform integration with Splunk®

This document describes how Splunk can be configured to collect alert log and audit log events on the NAS server platform.

MK-92HNAS082-01

May 2023

Table of Contents

Preface

About this document

This document describes how Splunk can be configured to collect alert log and audit log events on the NAS server platform.

Document conventions

This document uses the following typographic convention:

Convention	Description
Bold	• Indicates text in a window, including window titles, menus, menu options, buttons, fields, and labels. Example: **Click OK**. • Indicates emphasized words in list items.
Italic	Indicates a document title or emphasized words in text.
`Monospace`	Indicates text that is displayed on screen or entered by the user. Example: `pairdisplay -g oradb`

Intended audience

This document is intended for customers and Hitachi Vantara partners who license and use NAS Platform.

Accessing product downloads

Product software, drivers, and firmware downloads are available on Hitachi Vantara Support Connect: https://support.hitachivantara.com/.

Log in and select Product Downloads to access the most current downloads, including updates that may have been made after the release of the product.

Comments

Please send us your comments on this document to doc.comments@hitachivantara.com. Include the document title and number, including the revision level (for example, -07), and refer to specific sections and paragraphs whenever possible. All comments become the property of Hitachi Vantara.

Getting Help

Hitachi Vantara Support Connect is the destination for technical support of products and solutions sold by Hitachi Vantara. To contact technical support, log on to Hitachi Vantara Support Connect for contact information: https://support.hitachivantara.com/en_us/contact-us.html.

Hitachi Vantara Community is a global online community for customers, partners, independent software vendors, employees, and prospects. It is the destination to get answers, discover insights, and make connections. **Join the conversation today!** Go to community.hitachivantara.com, register, and complete your profile.

Chapter 1: Feature overview

The NAS server platform (Hitachi NAS platform and NAS module) can be integrated with Splunk® (https://www.splunk.com/). Splunk can be configured to collect alert log and audit log events, in addition to the ability to gather statistics about the NAS server system performance.

All instructions in this document assume that Splunk is already installed and working, and that all configuration, unless specified, is carried out using the Splunk Web interface.

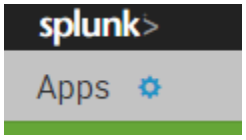
Chapter 2: Log collection

There are three steps to follow to configure Splunk to collect NAS server logs:

- Splunk additional setup
- Configuring Audit event input
- Configuring Alert input

Splunk additional setup

To enable NAS server logs to be collected by Splunk:

1. Install the CEFUtils Add-on on the Splunk server.
2. Download the CEFUtil Add-on from https://splunkbase.splunk.com/app/487/ and save locally.It does not need to be extracted before uploading.
3. Log in as Administrator to the Splunk Web interface, and from the home page, click the small blue wheel, which is at the top left of the home page.
4. 

5. Select “Install app from file”, and browse to the location where the CEFUtils Add-on is saved locally.
6. Once the Add-on is selected, press “Upload”. You will be prompted to restart Splunk to complete the update.

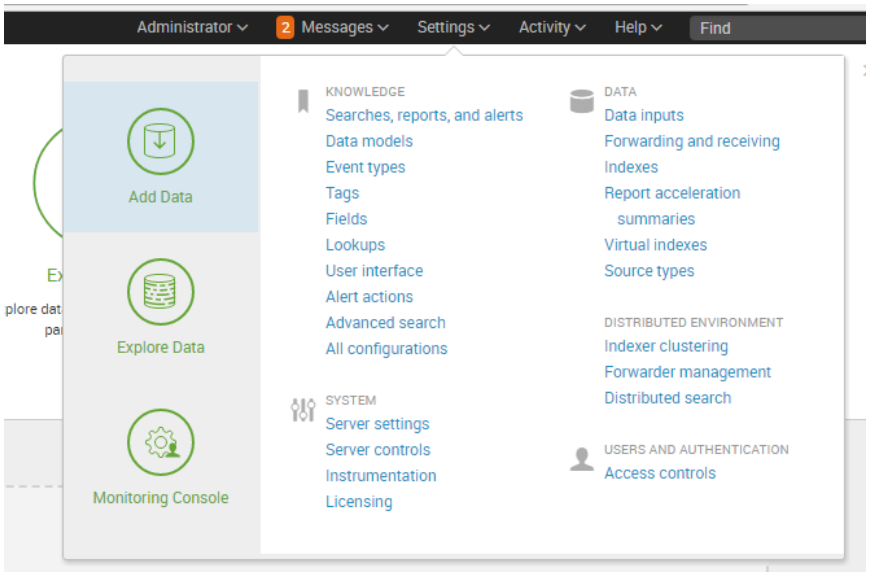
Once the Add-on is installed, Splunk needs to be configured with two separate inputs, which cover the standard syslog messages and the auditing events.

If the Splunk server has a firewall setup, then the ports and protocols configured in the next steps need to be allowed through the firewall, otherwise the firewall will block the traffic.

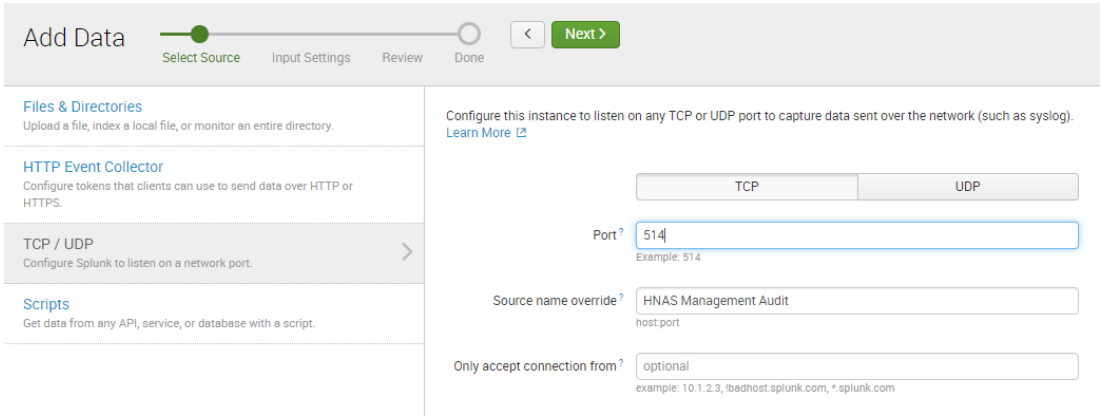
Configuring Audit Event input

To configure Splunk to accept NAS audit events:

1. On the Splunk home page, from the “Settings” menu on the top right side of the page, select “Data inputs”.

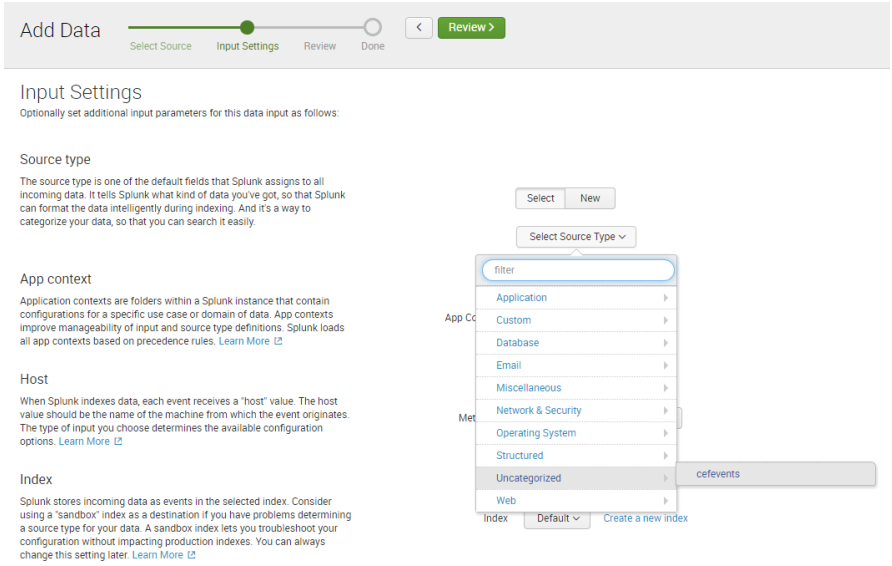


2. Select “Add new” from the “TCP” row in the “Local inputs” section. The following page appears:



3. Enter the TCP port - 514 is the default for Syslog, but if this is already in use, select another port and then change the value on the NAS server later.

4. Using the “Select Source Type” drop down button, select “Uncategorized” and then “cefevents”.

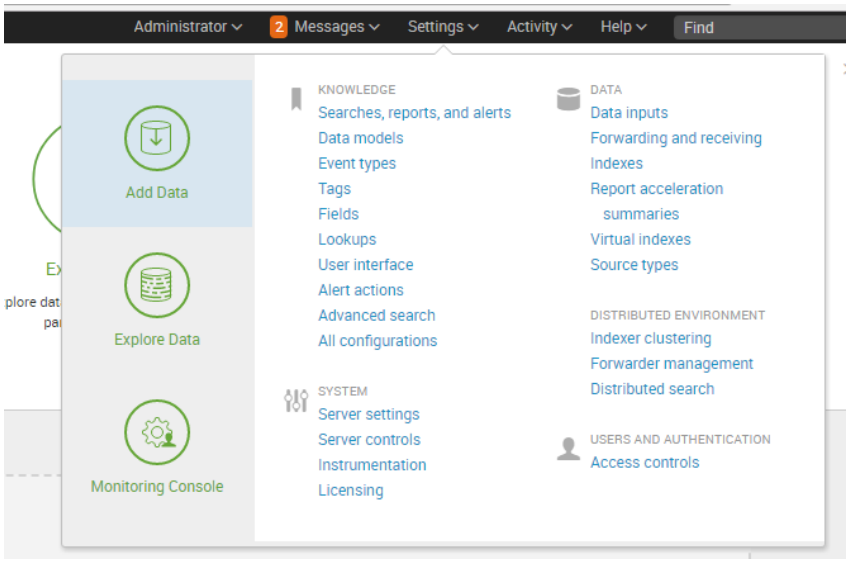


Once the Input Settings are completed, review the configured details, and submit them. Splunk now listens for TCP connection on the configured port, and when it received messages, it uses the CEFUtils Add-on to break them up according to the CEF fields.

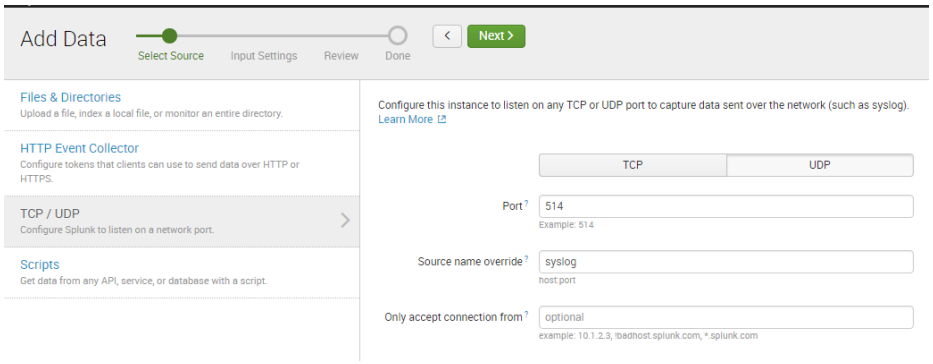
Configuring Alert input

To configure Splunk to accept NAS alerts:

1. On the Splunk home page, from the “Settings” menu on the top right side of the page, select “Data inputs”.



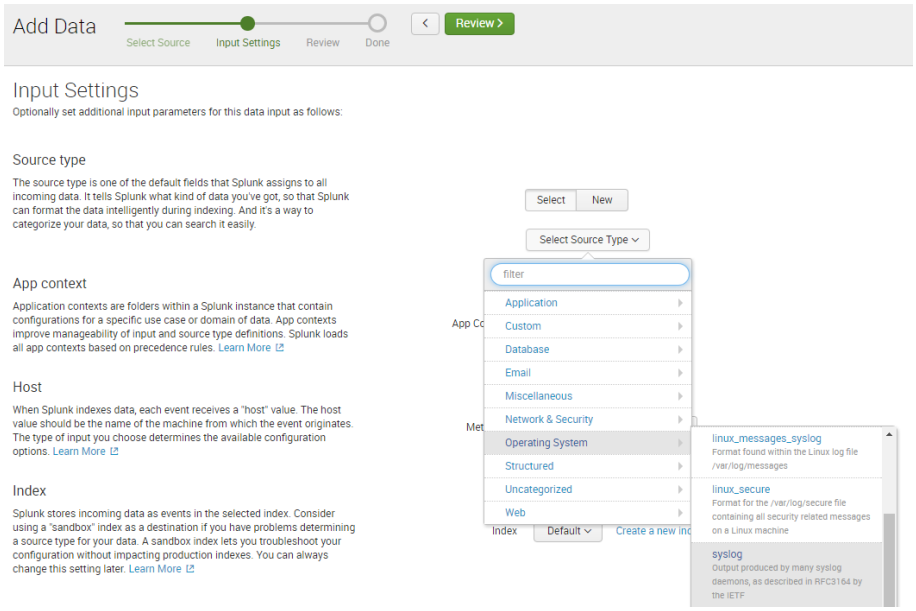
2. Check that UDP port 514 is not already configured to receive Syslog messages and if it’s not, select “Add new” from the “UDP” row in the “Local inputs” section. The following page appears:



3. If Syslog message collection is already configured, skip this section, and go to the NAS server setup section.

4. Enter the UDP port - 514 is the default for Syslog and cannot currently be changed on the NAS server, so it must be used.

5. Using the “Select Source Type” drop down button, select “Operating System” and then “syslog”.



Once the Input Settings are completed, review the configured details, and submit them. Splunk now listens for UDP messages on port 514, and when it received a message, it assumes that it is in syslog format, and decodes it accordingly.

Chapter 3: NAS server setup

The following settings now need to be applied to the NAS server so that it can send its alert messages to the Splunk server. The following instructions use Command Line Interface (CLI) commands, but the same functionality can also be achieved using the NAS Manager.

Syslog alert setup

Now configure the Syslog event frequency and destination. The example below configures the NAS server to send all alerts, at all levels to the Splunk server – this is the recommended setting. Although fewer events can be configured to be sent out, it may prove less useful during analysis later.

```
alert-syslog set –s i –w i –i i

alert-syslog add <SplunkServerAddress>
```

There is also a variable to make sure that the syslog alerts come from the Admin IP address – this is more useful for single node setups, but can help with clustered setups too:

```
set syslog_and_snmptrap_come_from_admin_vnode true
```

Management auditing setup

The management audit logs messages need to be configured separately, using the following command:

Note: if a port other than 514 was used in the Splunk setup, then it should be specified here, otherwise the `--port` option can be omitted from the command line.

```
audit-mgmt-log-server-add <SplunkServerAddress> [--port <port>]
```

Testing the setup

Once the Splunk server has been configured and the NAS server set up to send event and audit logs to the Splunk server, test the setup to make sure all types of event are sent to Splunk.
Generate some audit events:

- Login to the NAS server either through the NAS Manager or SSC – this action generates various authentication events.

Generate a test event log entry using the following CLI command:

- `alert-send-test`

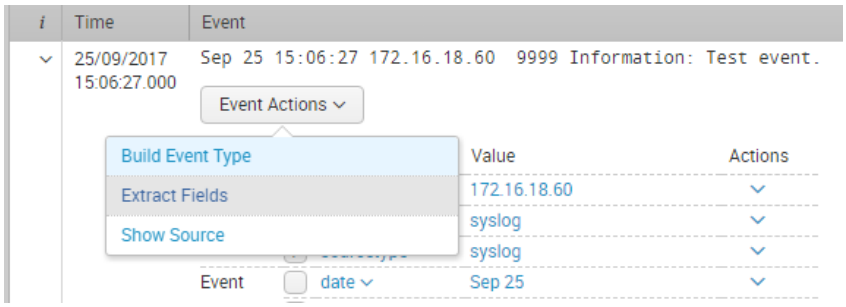
Below is an example of a Syslog test event, and a Logout auditing event from the SMU:

Chapter 4: Improving NAS Syslog messages in Splunk

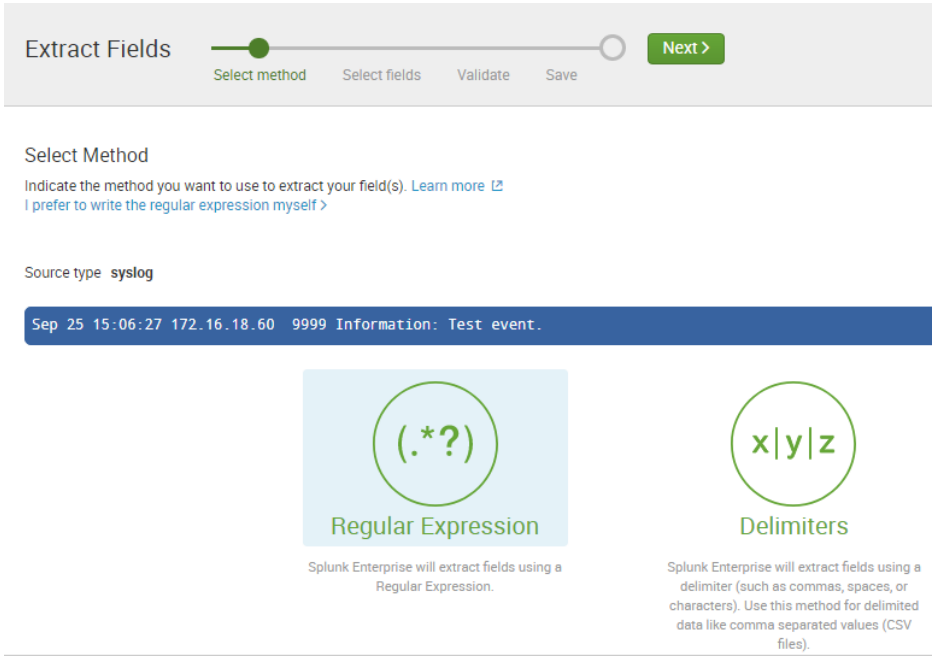
The NAS specific event log format can be extracted by Splunk to allow filtering, for example, on event severity, ID and other event parameters.
To filter the log:

1. Create a new search on the Splunk Web interface that includes an event log message from a NAS server. The test event created earlier would be suitable, for example.

2. Select the event log message, and expand it with the arrow on the left-hand side. Once expanded, select “Extract Fields” from the “Event Actions” button.



3. On the newly displayed page, select the “Regular Expression” option, and click **Next**.

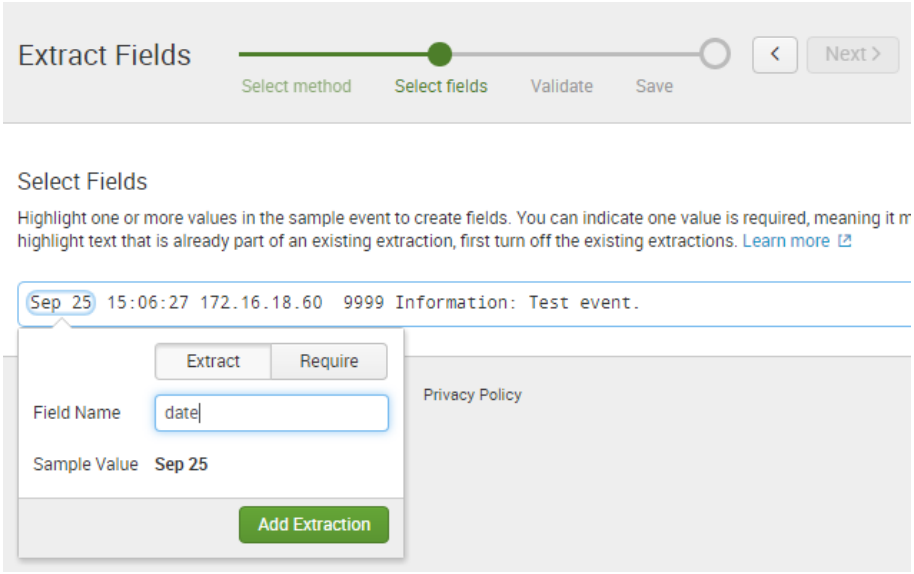


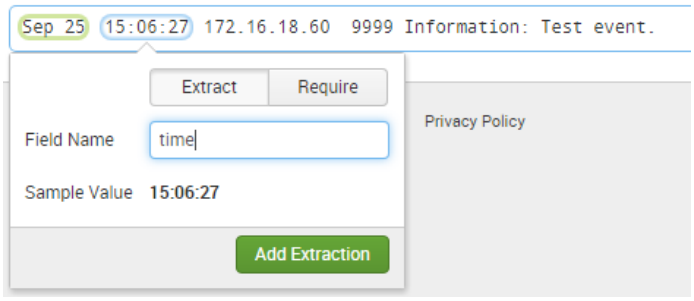
The full event text can now be broken down using regular expressions. These can be generated by selecting appropriate parts of the event message, and assigning field names to them. If Syslog messages from other systems are also being collected by Splunk, they will not fall within the regular expressions, so should pass without being filtered.

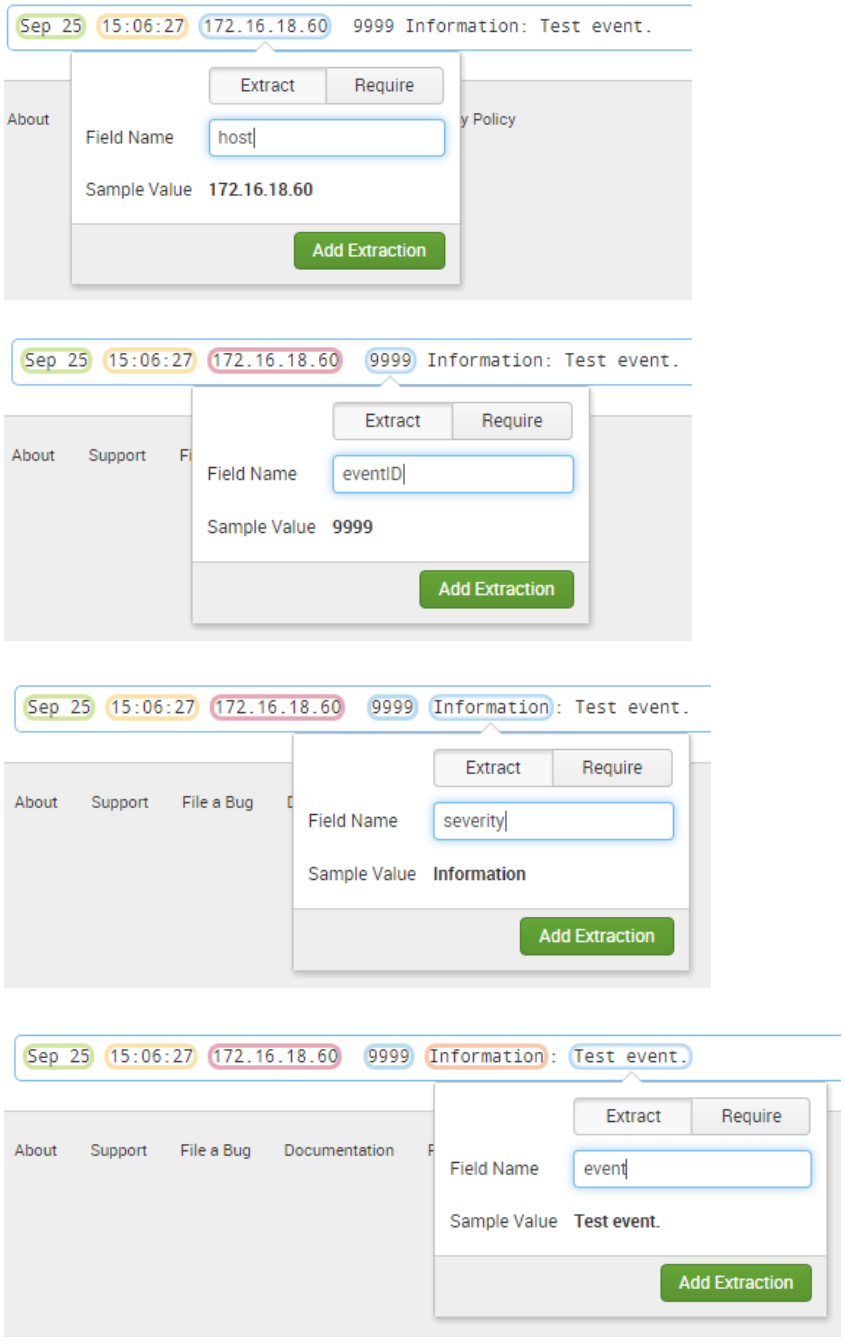
By selecting each part of the message in turn, it's possible to extract the following message specific information:

- date
- time
- host
- eventID
- severity *(exclude the : from the extraction)*
- event

4. Follow the instructions on the web page, and use the mouse to select extractions for each field named above. Refer to the screen shots below to show how to highlight each field:





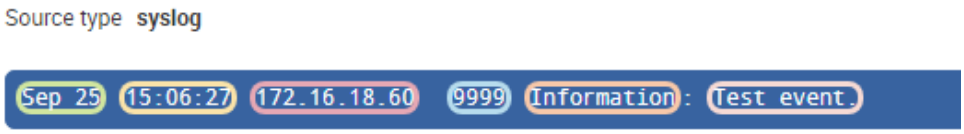


Once all fields have been defined and the regular expressions saved, the new filters will be applied to a selection of received messages, giving the option to edit/refine the filter if needed.

Optionally, name the extract and save it with a meaningful name. The default name will be:

syslog : EXTRACT-date,time,host,eventID,severity,event

The filter will be applied to all new Syslog messages, but should not impact non-NAS server Syslog messages on the same port. The purpose of extracting the fields is to allow them to be used in filters.



Limitations/Considerations

The following limitations need to be considered:

- Each cluster node needs to have a public management address to be able to communicate with the Splunk server.
- Syslog events can be lost when the NAS server system reboots, due to network unavailability. They are also sent via UDP, so onward transmission by the network is not guaranteed.
- When using a cluster, the events are not necessarily sent from the same address, even with the `syslog_and_snmptrap_come_from_admin_vnode` variable set. The admin EVS sends most events, but can select another serving EVS address if there is no route from the admin EVS to the Splunk server.
- Auditing events come from both physical node addresses, rather than specific admin address, as they are specific to the individual node.

Note: The use of multiple addresses makes filtering for events from a complete system more difficult, as they may have come from any of the configured IP addresses.

Chapter 5: Splunk NAS Add-on

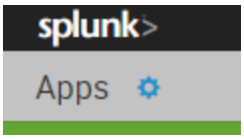
The purpose of this Add-on is to allow performance statistics to be collected from NAS servers. Different statistics may be available in different NAS software versions, but all versions should work.

Once the statistics are ingested into Splunk, they can then be manipulated in a variety of ways. The Add-on does not make any attempt to provide graphs or dashboards with the ingested statistics, these will need to be constructed separately, on a per-use case basis.

Installing the Add-on

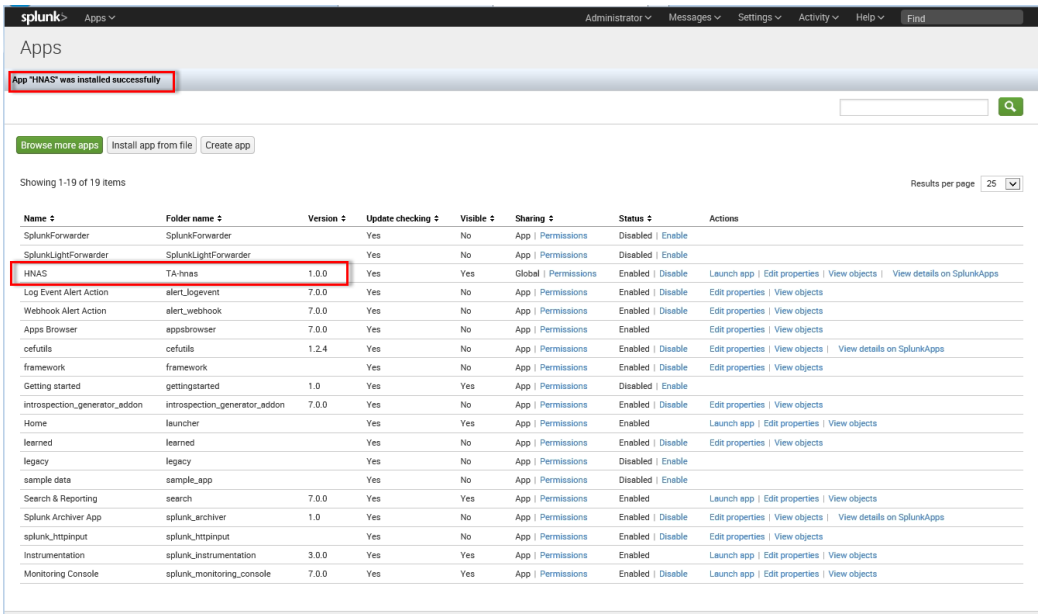
The Add-on is provided as a tgz (compressed) file, which needs to be uploaded directly to Splunk. It does not need to be extracted before uploading.

Log in as an administrator to the Splunk Web interface, and from the Splunk Home page, click the small blue wheel, which is at the top left of the home page.

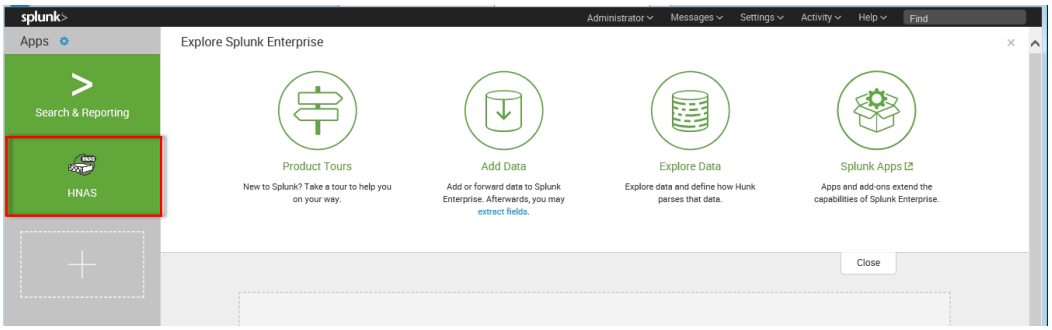


Select “Install app from file”, and browse to the location where the Add-on file is stored on your local hard drive. Once the Add-on file is selected, press “Upload”. You are prompted to restart Splunk to complete the update.

Once restarted, you should see confirmation as follows:

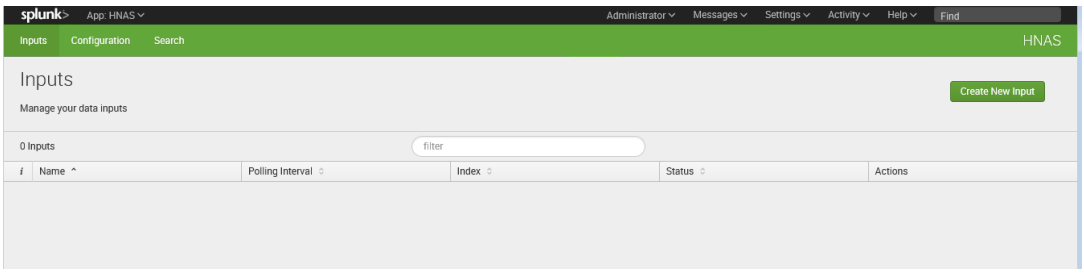


Once installed, the following new button should be present on the Splunk Home page as follows:



Adding NAS servers

Select the “HNAS” button from the Splunk Home page – a new page is displayed. This initially appears to be empty, until any NAS servers are configured for monitoring:



Once the new page loads, select the “Create New Input” button. The “Add HNAS-Stats” dialog appears.

Enter the following information:

- **Name** – A unique name for the server to be monitored. Note the character restrictions: Input Name must start with a letter and followed by alphabetic letters, digits or underscores. Keep in mind that you must do this for each node, so use the name of the node (for example ”arena_1”) not the cluster name (“arena”).

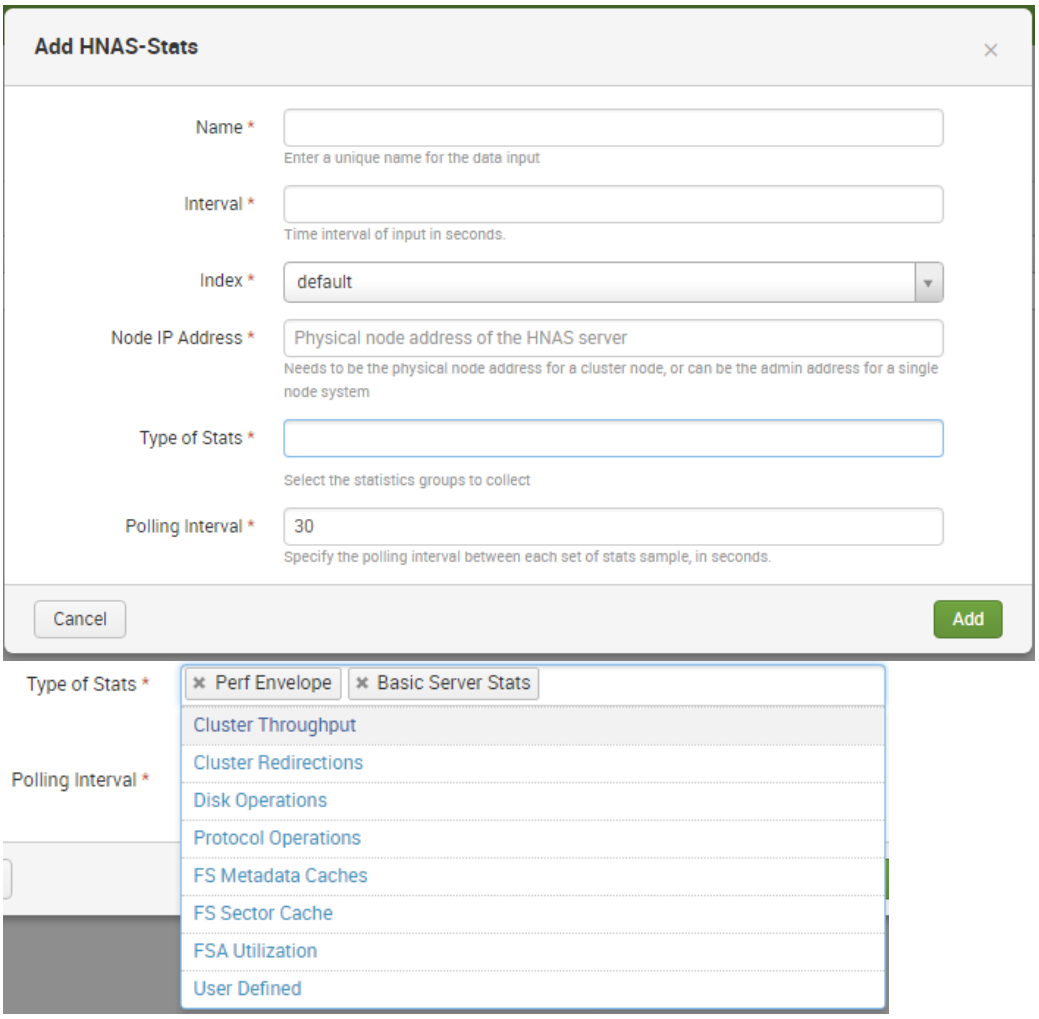
- **Interval** – This value must be set to 0. It represents the restart interval that Splunk uses for each Add-on. The value of 0 indicates that the process will restart automatically if it exits.

- **Index** – the Splunk index to which the statistics will be added – select this value from the drop-down list. Select “default” if unsure.

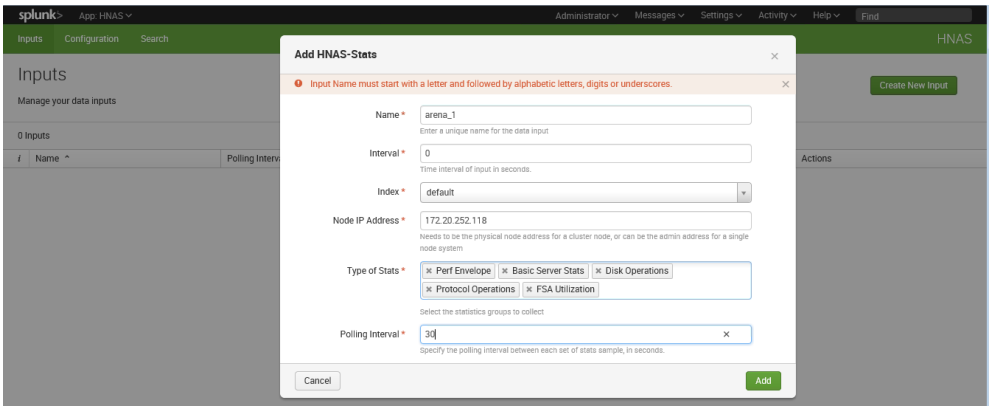
- **Node IP Address** – This needs to be an address of the physical node where stats are to be gathered from. If this system consists of only one node, then this can be the Admin EVS address. For multi-node clusters, each node needs to be added separately for monitoring, otherwise the statistics gathered will only relate to one of the nodes.

- **Type of Stats** – This is a multiple selection field, and allows various pre-defined statistic groups to be selected. We recommend at least Basic Server Stats. See below for more details.

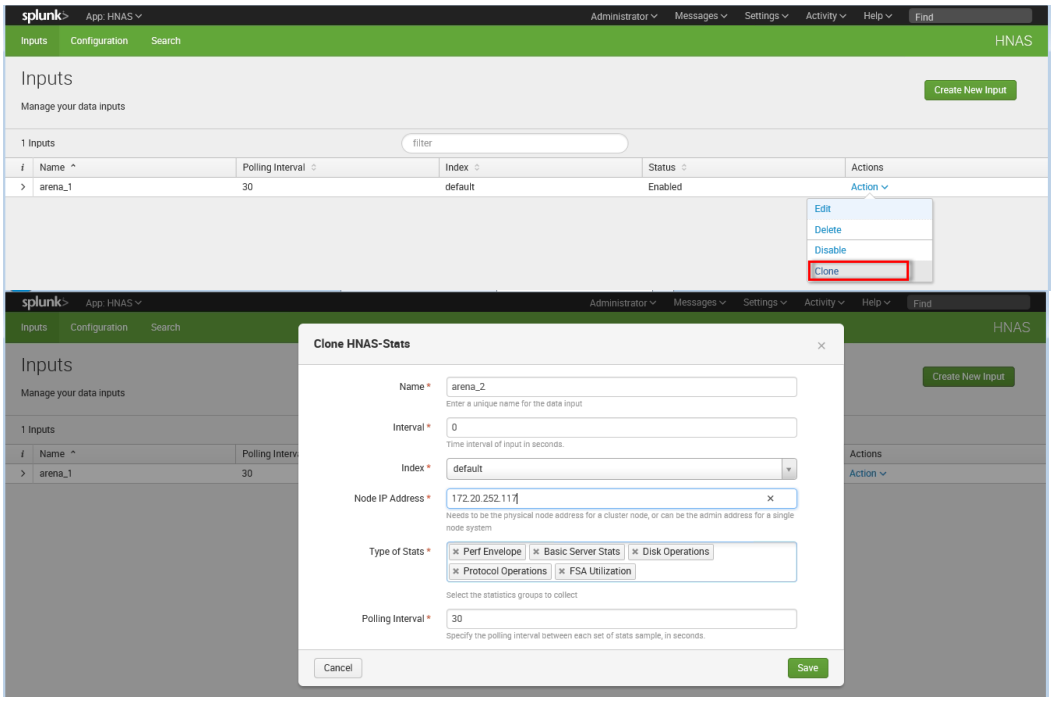
- **Polling Interval** – This is the delay, in seconds, between each statistic sample being collected. The default is 30 seconds. It is best not to go any lower than 30 seconds.



Example:

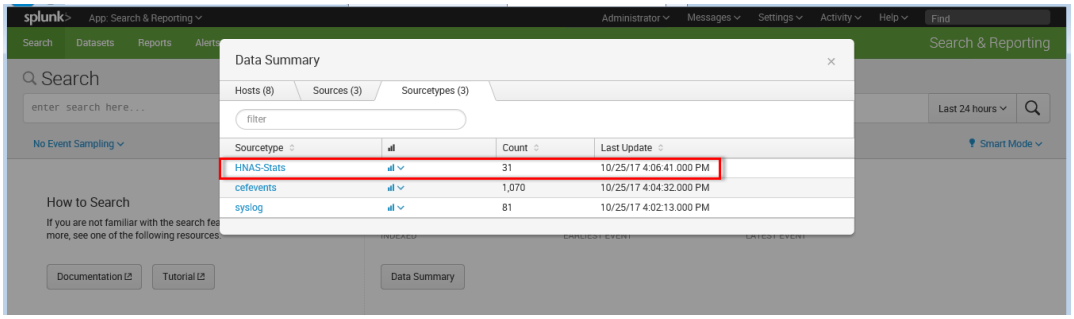


Once the server is added, you can use the menu in the Actions column on the right to clone and configure another node or server.

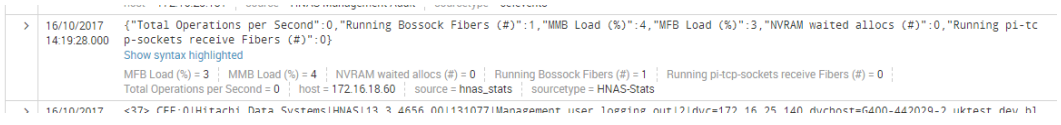


Examining statistics

Navigate to Splunk > Search > Data Summary to browse the new data generated. This can be seen in the “Sourcetype” column:

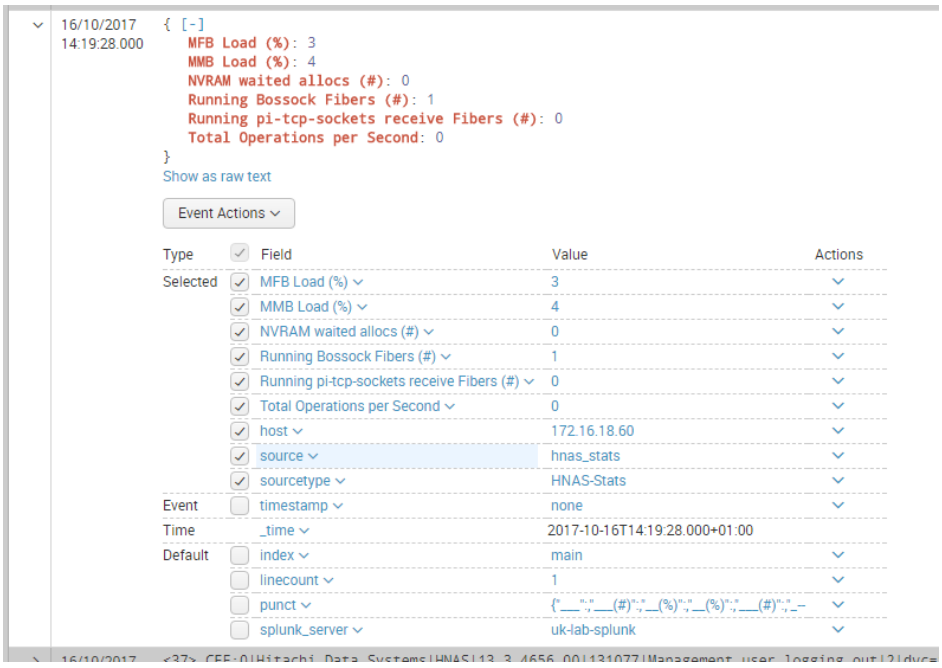


The gathered statistics are added to the Splunk index in JSON format – Statistic Name, followed by its value – below shows the “raw text” representation of an instance of the “Perf Envelope” statistics added to the Splunk index:



If you click on “Show syntax highlighted” it will show the same data but with JSON syntax high lighting:

And the same data again, but expanded to allow individual fields to be selected for actions and events:



Predefined statistics groups

The predefined statistic groups are based on the ones already available using the Rusc tool, which is available on the SMU.

Note: if a statistic is defined in multiple selected groups, only one instance of that statistic will be added to Splunk.

Perf Envelope:

["Total Operations per Second", "MMB Load (%)","MFB Load (%)", "NVRAM waited allocs (#)", "Running Bossock Fibers (#)", "Running pi-tcp-sockets receive Fibers (#)"]

Basic Server Stats:

["Total Operations per Second", "MMB Load (%)", "MFB Load (%)", "Ethernet Throughput RX (Mbps)", "Ethernet Throughput TX (Mbps)", "FibreChannel Throughput RX (Mbps)", "FibreChannel Throughput TX (Mbps)", "NVRAM waited allocs (#)"]

Cluster Throughput:

["HSSI-C1 Throughput RX (Mbps)", "HSSI-C1 Throughput TX (Mbps)", "HSSI-C2 Throughput RX (Mbps)", "HSSI-C2 Throughput TX (Mbps)"]

Cluster Redirections:

["Total Operations Received per Second", "Total Forwarded Operations Received per Second", "Total Operations Forwarded per Second"]

Disk Operations:

["Current Virtual Disk Requests (#)", "Disk Read Latency (ms)", "Disk Write Latency (ms)", "Disk Stripe Write Latency (ms)"]

Protocol Operations:

["Protocol : iSCSI Operations per Second", "Protocol : NFS Operations per Second", "Protocol : SMB Operations per Second", "Protocol : SMB2 Operations per Second"]

FS Metadata Caches:

["wfile cache hits (%)", "wdir cache hits (%)", "wtree cache hits (%)", "obj_store root onode cache hits (%)", "obj_store non-ind-obj leaf onode cache hits (%)", "obj_store ind-obj leaf onode cache hits (%)"]

FS Sector Cache:

["Sector Cache Read Hits (%)", "Sector Cache ReadAhead Hits (%)", "Sector Cache Write Hits (%)"]

FSA Utilization:

["FSA Cache usage (%)", "Heap Usage (%)", "Running Bossock Fibers"]

User Defined:

User defined statistics are covered in the next section.

Gathering Additional Statistics – User Defined Statistics

In certain situations, it may be desirable to monitor other statistics, and the stats available may differ as server builds change, or it may be necessary to monitor statistics specific to a certain filesystem for example.

With that requirement in mind, the "User Defined" statistics group has been added. By default, it has only a single value associated with it. Over 5000 statistics are available from the NAS server, so it's not practical, or desirable to monitor them all, especially as Splunk licensing is based on the volume of data ingested per month, and the more statistics that are monitored, the more Splunk licensed bandwidth they use up.

Getting a List of Supported Statistics

As the statistics are version specific, and depend on various configuration options, it's not possible to supply a list here, but it is possible to ask the NAS server what statistics it supports as follows:
Telnet to the NAS server admin address, connect to TCP port 11106, and type LIST. The server lists all the currently available statistics. End the connection with QUIT. See the truncated example below:

```
root@uk-lab-splunk:/opt/splunk/etc/apps/TA-hnas/bin# telnet 172.16.18.60 11106
Trying 172.16.18.60...
Connected to 172.16.18.60.
Escape character is '^]'.
LIST
0       Number of ssfs operations executed directly by the BOS Thread
1       sockets: New connections while under msgb stress
2       sockets: New connections denied while under msgb stress
3       sockets: New connections while under heap stress
4       sockets: New connections denied while under heap stress
....
6612    Reverse migration: wait for pool slot timed out (#)
6613    Reverse migration: wait for pool slot didn't time out (#)
.
QUIT
Connection closed by foreign host.
```

The statistics supported are prefixed by a number – the text on the right-hand side is what is required to build a User Defined statistics group.
Currently, the only way to update the group is to directly edit the Python file, which is part of the Add-on source, and the User Defined statistics will then be the same for each system. It is not possible to generate a different User Defined statistics group for each system monitored.

The file that needs to be edited is statsdef.py, which on Linux should be in the /opt/splunk/etc/apps/TA-hnas/bin/ folder and on Windows should be in the C:\Program Files\Splunk\etc\apps\TA-hnas\bin\ folder. The bottom line of the file is the only one that should be edited, and must maintain the format:

```
HNASstats["User Defined"] = ["Full name of stat 1", "Full name of stat 2",
"Full name of stat 3"]
```

By default, the User Defined statistics groups is as below, which is provided as an example:

```
HNASstats["User Defined"] = ["Protocol : FTP Operations per Second"]
```

Hitachi Vantara

Corporate Headquarters 2535 Augustine Drive
Santa Clara, CA 95054 USA www.HitachiVantara.com community.HitachiVantara.com

Regional Contact Information
Americas: +1 866 374 5822 or info@hitachivantara.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hitachivantara.com
Asia Pacific: +852 3189 7900 or info.marketing.apac@hitachivantara.com