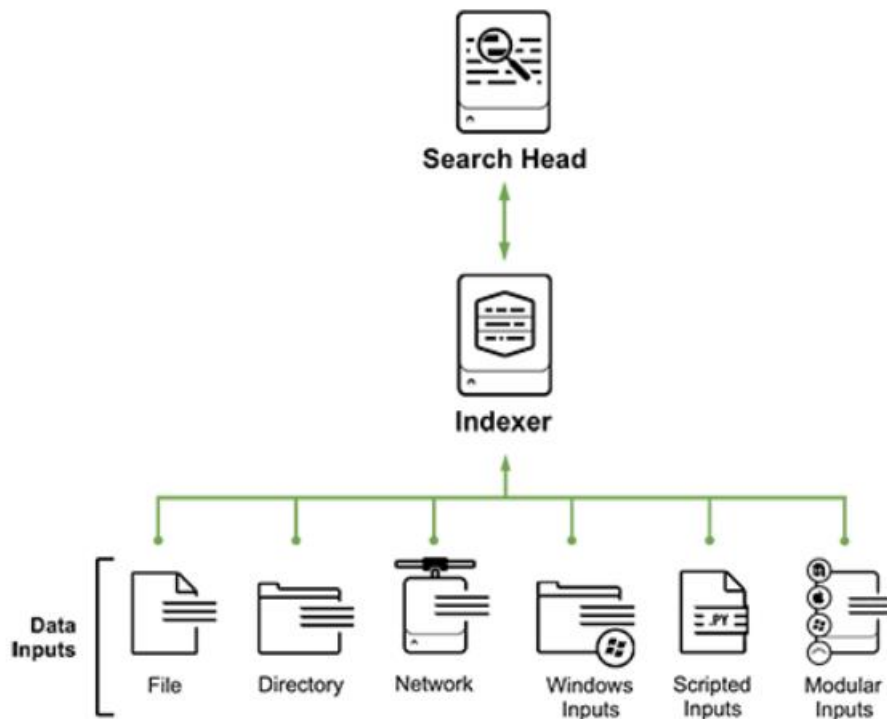


С311. Splunk — SPL запросы и визуализация логов

Загрузка данных в систему



В системе можно выделить 5 основных источников сбора логов (это не полный список):

Files and Directories: Splunk может разово забирать или мониторить конкретный файл или директорию с файлами, причем самостоятельно следит за изменением

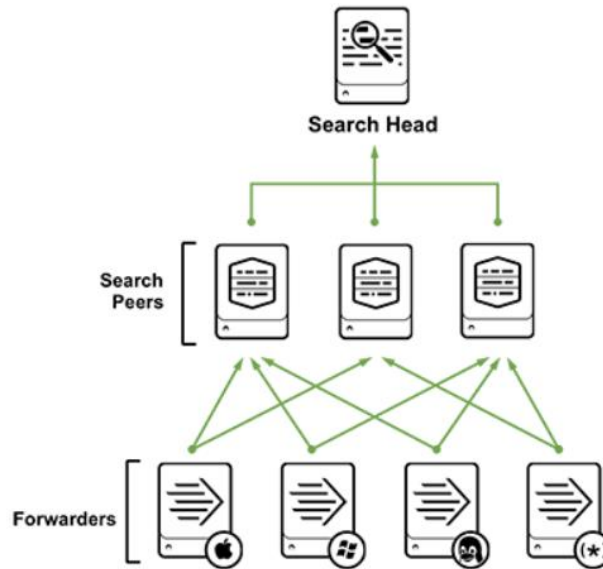
Network events: данные, поступающие с сетевых портов (syslog например)

Windows sources: журналы событий (event log) Windows, события AD

Scripted Inputs: данные собираемые посредством скриптов

Modular Inputs: забор данных из специфических платформ, систем и приложений

В данной работе, для наглядности мы будем использовать наиболее простой метод. Мы просто загрузим тестовый файл в Splunk с локального компьютера. Понятно, что в истории с Enterprise использованием так никто не делает, и как раз используются варианты описанные выше вместе с агентами (forwarder), стоящими на целевых системах, и тогда инфраструктура выглядит следующим образом:



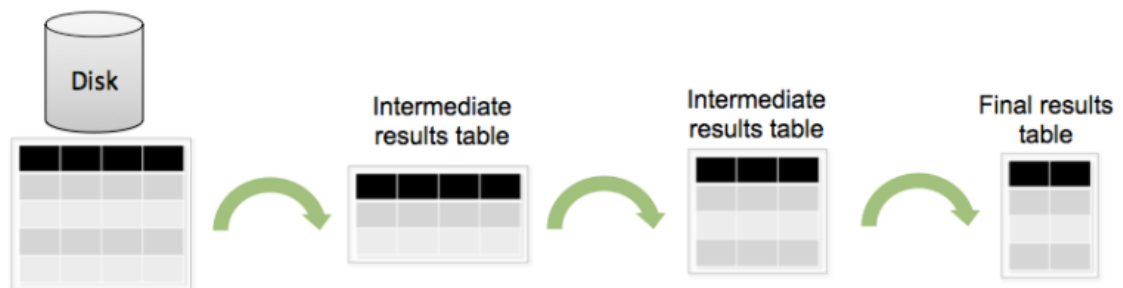
Но в нашем учебном примере, нам будет достаточно одного скаченного на локальный компьютер бесплатного [Splunk Enterprise Free](#).

Теперь когда Вы скачали данные и установили Splunk их надо загрузить в него. На самом деле это достаточно просто ([инструкция](#)), потому что данные заранее подготовлены. **Важно!** Не нужно разархивировать архив.

SPL запросы

Ключевые особенности языка SPL:

- 140+ поисковых команд
- Синтаксис похож на Unix pipeline и SQL и оптимизирован на **данные с временной отметкой**
- SPL позволяет искать, фильтровать, модифицировать, обогащать, совмещать и удалять
- SPL включает функционал **машинного обучения** и **поиска аномалий**



Структура SPL:

Стандартно SPL запрос можно разделить на несколько этапов: фильтрация и выбор нужных данных,

затем создание новых полей на основе уже существующих, затем агрегирование данных и вычисление статистик, и в конце переименование полей, сортировка другими словами, обогащение вывода.

search and filter | munge | report | cleanup

sourcetype=access*

| eval KB=bytes/1024

| stats sum(KB) dc(clientip)

| rename sum(KB) AS «Total KB» dc(clientip) AS «Unique Clients»

После того как Вы загрузили данные в систему вы можете осуществлять поиски по ним (ниже примеры запросов, с результатами выполнения):

Поисковый интерфейс имеет следующий вид:

Кнопка запуска поиска

Поисковая строка, используется т.н. SPL

Кнопка выбора времени

Визуализация событий во времени

Поля извлеченные из данных

Найденные события

#	Time	Event
>	3/31/16 4:22:04.730 AM	2016-03-31T09:22:04.730Z,33.9815,-117.1323333,19.61,0.82,m1,21.65,0.09249,0.22,ci,ci37325015,2016-03-31T09:26:25.524Z,"7km W of Calimesa, CA",earthquake,0.44,0.98,0.145,21,automatic,ci,ci sourcetype = eq
>	3/31/16 4:20:33.250 AM	2016-03-31T09:20:33.250Z,38.8296661,-122.8468323,1.37,0.68,md,8,178,0.0113,0.01,nc,nc72615215,2016-03-31T09:22:13.864Z,"9km NW of The Geysers, California",earthquake,0.44,0.79,0.11,2,automatic,nc,nc sourcetype = eq
>	3/31/16 4:20:04.360 AM	2016-03-31T09:20:04.360Z,38.8373337,-122.8079987,1.59,0.18,md,5,136,0.01401,0,nc,nc72615210,2016-03-31T09:21:44.084Z,"7km MNW of Cobb, California",earthquake,0.56,0.98,0.11,2,automatic,nc,nc sourcetype = eq

Поиск и фильтрация:

В Splunk можно «как в гугле» искать события по ключевому слову, или набору ключевых слов разделенных стандартными логическими операторами, примеры ниже. Также вы можете в любой момент актуализировать свой поиск, выбрав нужный вам временной интервал как в меню справа, так и центральной зеленой гистограмме, которая показывает количество событий в определенный период

времени.

Поиск по ключевому слову: sourcetype=access* http

The screenshot shows a Splunk search interface with the query 'sourcetype=access* http'. The search results are displayed in a table with columns for Time and Event. The table shows several events from March 30, 2016, with details such as IP addresses, user agents, and hostnames. The interface includes a search bar, a timeline visualization, and various filters and options.

Фильтрация: sourcetype=access* http clientip=87.194.216.51

The screenshot shows a Splunk search interface with the query 'sourcetype=access* http clientip=87.194.216.51'. The search results are displayed in a table with columns for Time and Event. The table shows several events from March 30, 2016, with details such as IP addresses, user agents, and hostnames. The interface includes a search bar, a timeline visualization, and various filters and options.

Комбинация: sourcetype=access* http clientip=87.194.216.51 (503 OR 504)

The screenshot shows a Splunk search interface with the query 'sourcetype=access* http clientip=87.194.216.51 (503 OR 504)'. The search results are displayed in a table with columns for Time and Event. The table shows several events from March 30, 2016, with details such as IP addresses, user agents, and hostnames. The interface includes a search bar, a timeline visualization, and various filters and options.

[Splunk. Введение в анализ машинных данных. Примеры SPL запросов и визуализация логов](#)