

Creación de una cuenta elastic

Permite el despliegue de entorno mediante el servicio Elasticsearch Service.

<https://www.elastic.co/elasticsearch/service>

Corresponde a una implementación completamente cargada en la nube en uno de los tres proveedores disponibles AWS, Microsoft Azure o Google Cloud.

Elasticsearch Service Create deployment				
Deployment name	Status	Version	Cloud region	Manage deployment
electivo_stack_elk	Healthy	8.5.0	GCP - Iowa (us-central1)	

Propósito de uso de la pila de elastic

El objetivo es analizar los logs generados por el servidor web **nginx** encargado de servir una página web creada especialmente para esta instancia.

Acerca del servidor para el levantamiento de la aplicación web

Se ha hecho el levantamiento de un servidor virtual (virtual machine) en la nube de Google. Con el propósito de alojar en este una pagina web.

La dirección ip del servidor es 34.125.62.110.

Las instancias de VM son máquinas virtuales altamente configurables para ejecutar cargas de trabajo en la infraestructura de Google. [Más información](#)

Filtro Ingresar el nombre o el valor de la propiedad

<input type="checkbox"/>	Estado	Nombre ↑	Zona	Recomendaciones	En uso por	IP interna	IP externa	Conectar
<input type="checkbox"/>	✓	instance-1	us-west4-b	Ahorrar \$26/mes		10.182.0.12 (nic0)	34.125.62.110 (nic0)	SSH ▾

No es seguro | 34.125.62.110

Link_U Traducir Cambridge Diction... Recibidos (441) - hr... Leccion Adultos Adult Bible Study G... GitHub

Welcome to nginx!

If you see this page, the nginx web server is successfully installed and working. Further configuration is required.

For online documentation and support please refer to nginx.org.
Commercial support is available at nginx.com.

Thank you for using nginx.

Conexión remota entre elasticsearch service y el servidor de la aplicación web

El primer paso a realizar es la conexión entre los elementos involucrados. Es decir, debemos permitirle al servicio de elastic tener acceso a los logs generados por nginx en el servidor remoto.

Para ello, elasticService nos proporciona El agregado de integraciones.


Get started by adding integrations

To start working with your data, use one of our many ingest options. Collect data from an app or service, or upload a file. If you're not ready to use your own data, add a sample data set.

[+ Add integrations](#) [Try sample data](#) [Upload a file](#)

Elastic service proporciona una gran cantidad de integraciones para diferentes softwares disponibles en el mercado. Para nuestro caso, usaremos la integración especialmente diseñada para nginx.

[Back to integrations](#)



Nginx

Elastic Agent

Version
1.6.0

Agent policies
2

+ Add Nginx

Overview

Integration policies

Assets

Settings

API reference

Integration policy	Version	Agent policy	Last updated by	Last updated	Agents	Actions
nginx-1	v1.6.0	Agent policy 1 rev. 2	230757112	3 days ago	1	...

Rows per page: 20

< 1 >

Después de configurar la integración es necesario crear un agente en el servidor donde corre nginx. Es decir, En el servidor que sirve la aplicación web mencionada más arriba.

Nginx integration added

To complete this integration, add **Elastic Agent** to your hosts to collect data and send it to Elastic Stack

[Add Elastic Agent later](#) [Add Elastic Agent to your hosts](#)

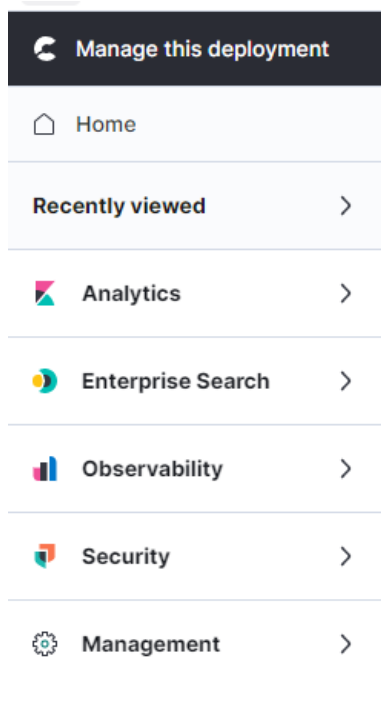
Para la instalación de dicho agente en la maquina remota se debe acceder a la misma y en la terminal se debe ejecutar una serie de comandos a continuación:

1. `curl -L -O https://artifacts.elastic.co/downloads/beats/elastic-agent/elastic-agent-8.5.0-amd64.deb`
2. `sudo dpkg -i elastic-agent-8.5.0-amd64.deb`
3. `sudo elastic-agent enroll --url=https://61fe17aa5c6b4a6781b66c83719e43d8.fleet.us-central1.gcp.cloud.es.io:443 --enrollment-token=dWVvVmQ0UUJrTmZINVh5NzBVZHk6cjQweTI0V2hUVzZNNDc2OGJMVG11UQ==`
4. `sudo systemctl enable elastic-agent`
5. `sudo systemctl start elastic-agent`

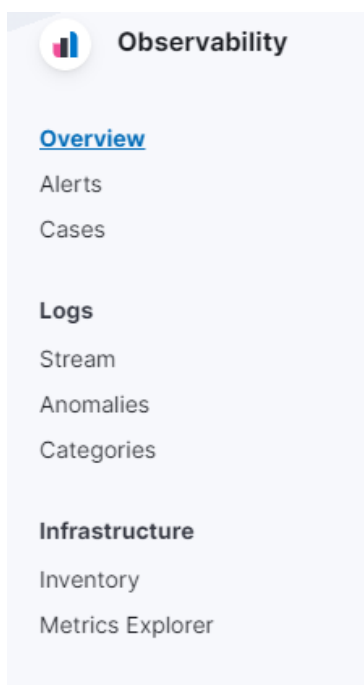
Lo que se hace aquí es lo siguiente:

- en la línea 1: se descarga en la maquina remota el agente de elastic.
- Línea 2: se instala en la maquina remota el agente.
- Se proporciona una clave especial que permita la vinculación del agente con el servicio de elastic.
- Líneas 4 y 5. Se activa e inicia el agente como un servicio del sistema operativo.

Ahora se chequea la conexión entre ambos componentes de software:



En la sección Observability acudimos a hora a la subsección log/stream



Ahora podemos en la pantalla que se nos están enviando mensajes del log de nginx

Stream

Customize Highlights

Last 1 day

Stream live

Nov 14, 2022	event.dataset	Message	
14:04:05.001	elastic_agent.metricbeat	[elastic_agent.metricbeat][info] Non-zero metrics in the last 30s	06 PM
14:04:05.463	elastic_agent.metricbeat	[elastic_agent.metricbeat][info] Non-zero metrics in the last 30s	
14:04:12.000	nginx.access	[nginx][access] 127.0.0.1 "GET /nginx_status? HTTP/1.1" 404 125	Mon 14
14:04:12.304	elastic_agent.metricbeat	[elastic_agent.metricbeat][error] Error fetching data for metrics et nginx.status: error fetching status: HTTP error 404 in : 404 Not Found	06 AM
14:04:22.000	nginx.access	[nginx][access] 127.0.0.1 "GET /nginx_status? HTTP/1.1" 404 125	12 PM
14:04:22.305	elastic_agent.metricbeat	[elastic_agent.metricbeat][error] Error fetching data for metrics et nginx.status: error fetching status: HTTP error 404 in : 404 Not Found	

```

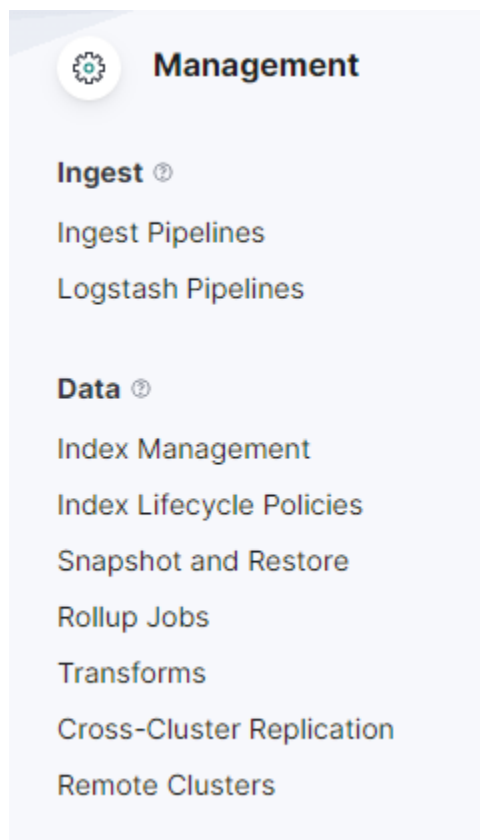
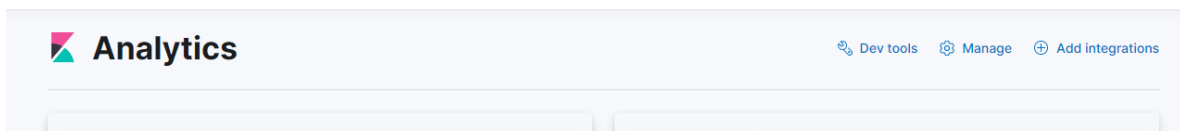
14:22:13.000 nginx.access [nginx][access] 152.174.41.155 "GET /? HTTP/1.1" 304 0
14:22:16.000 nginx.access [nginx][access] 152.174.41.155 "GET /? HTTP/1.1" 304 0
14:22:18.000 nginx.access [nginx][access] 152.174.41.155 "GET /? HTTP/1.1" 304 0
14:22:19.000 nginx.access [nginx][access] 152.174.41.155 "GET /? HTTP/1.1" 304 0
14:22:22.000 nginx.access [nginx][access] 127.0.0.1 "GET /nginx_status? HTTP/1.1" 404 125
14:22:22.304 elastic_agent.metricbeat [elastic_agent.metricbeat][error] Error fetching data for metrics
et nginx.status: error fetching status: HTTP error 404 in : 4
04 Not Found

```

Uso de los datos del log de nginx con ElasticService

Ahora estamos en condiciones de “jugar” con elasticsearch y kibana.

En esta sección primeramente indicaremos como acceder al índice que se ha generado al instalar el agente en el servidor remoto. Recordemos que este índice contiene información sobre el log de acceso y de errores del servidor web nginx instalado en la maquina remota.



Index Management

[Index Management docs](#)

Indices [Data Streams](#) Index Templates Component Templates

Data streams store time-series data across multiple indices. [Learn more.](#)

☐ Include stats

View 1

Search...

Reload

<input type="checkbox"/> logs-nginx.access-default	● green	1	
<input type="checkbox"/> logs-nginx.error-default	● green	1	

logs-nginx.access-default

Health

● green

Last updated

November 15th, 2022
10:36:02 AM

Storage size

11.1mb

Indices

1

Timestamp field

@timestamp

Generation

1

Index template

logs-nginx.access

Index lifecycle policy

logs

logs-nginx.error-default

Health

● green

Last updated

November 11th, 2022
5:38:12 PM

Storage size

49kb

Indices

1

Timestamp field

@timestamp

Generation

1

Index template

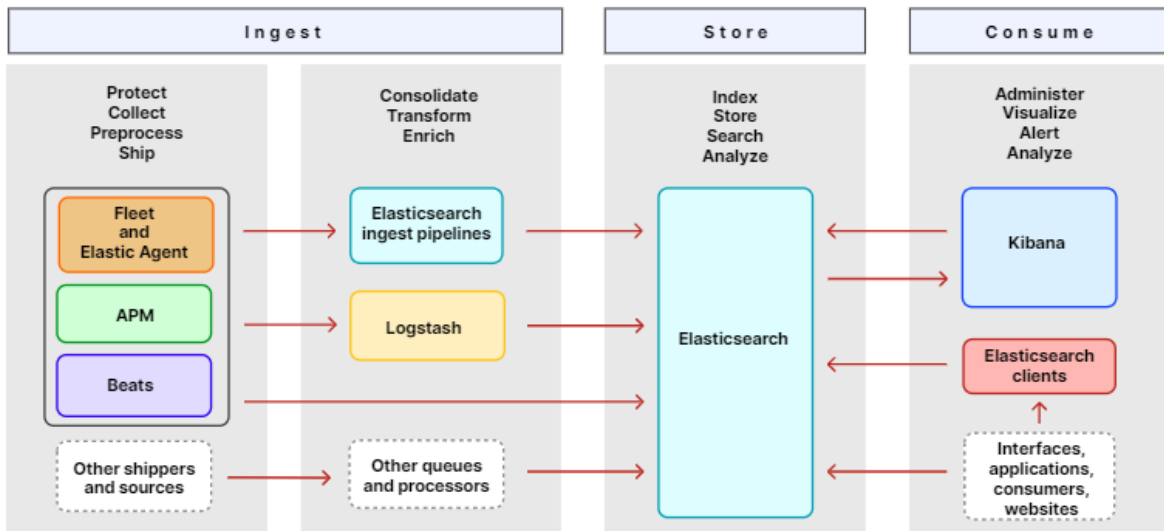
logs-nginx.error

Index lifecycle policy

logs

En las imágenes anteriores se ha presentado como localizar un índice alimentado remotamente.

Components of the Elastic Stack



<https://www.elastic.co/guide/en/welcome-to-elastic/current/stack-components.html>

¿Que es un data stream en elasticsearch?

Respuesta: <https://aravind.dev/elastic-data-stream/#:~:text=Data%20streams,-Logs%2C%20metrics%2C%20traces&text=It%20rolls%20over%20the%20index,entire%20hidden%20collection%20of%20indices.>

Sobre los casos de uso

Idea inicial

Para caso del ejemplo utilizaremos el índice de acceso de nginx para realizar consultas sobre el mismo.

Cual es la idea de lo que se hará:

Meta: detectar direcciones ip registradas en listas negras que intenten acceder a un determinado sitio web para posteriormente restringir su acceso.

Roadmap de la solución:

1. Se chequea en tiempo real las direcciones ip que intentan acceder a los recursos del sitio web.
 - a. Se debe ir revisando cada cierto tiempo los últimos registros del log de acceso del servidor web. La ventana de tiempo puede ser de 30 segundos.
 - i. Chequear log de acceso cada 30 segundos.
 - ii. El ultimo log de la ventana de tiempo anterior debe quedar marcado para que sirva como inicio de la nueva ventana.
 - iii. Si la dirección ip se repite en el lapso de la ventana de tiempo, esta no debe volver a chequearse.

Idea Final

Aplicación Web: proporcionar monitoreo personalizado de servicios web mediante el servicio infoweb. Esto le permitirá al administrador del servicio que desea monitorear tener información personalizada acerca del comportamiento del mismo y sin la necesidad de tener que hacer uso directo del stack (con una curva de aprendizaje no menor)

Meta: Proporcionar información a la entidad encargada de un sitio web respecto de la actividad de los visitantes de la misma como así también información sobre la maquina en la cual se ejecuta la aplicación web.

- Alcance:
 - El prototipo cuenta solamente con un usuario activo el cual tiene una aplicación web conectada a la cual se le hará el monitoreo.
 - El prototipo asume que la aplicación web ya está conectada al servicio. (no hay forma de conectar una aplicación en esta etapa).
- Funcionalidades:
 - Se obtendrá la cantidad de visitas por mes.
 - Como se hará: Se necesitan dos marcas de tiempo las cuales definan un intervalo en el índice de elasticsearch sobre el cual se realizará la búsqueda de las direcciones ip.
 - Las ip repetidas se contarán solamente una vez.

Puntos importantes a considerar

- Log original generado por nginx: pendiente de ver
- De qué manera se parsean los datos desde el log hacia elasticsearch.
- Como ver el formato de los datos entregados.
- Como acceder al índice con el contenido del log de nginx desde la aplicación mediante el uso de la librería de elasticsearch.
- Confirmar si se trata de un index o de un flujo de datos.
- Nombre del index: .ds-logs-nginx.access-default-2022.11.11-000001

```
1 {
2   "_index": ".ds-logs-nginx.access-default-2022.11.11-000001",
3   "_id": "3031e408kNFH5xy7KiZU",
4   "_version": 1,
5   "_score": 0,
6   "_source": { ... },
141 },
142   "_fields": { ... },
345 }
346 }
```

```
1 {
2   "_index": ".ds-logs-nginx.access-default-2022.11.11-000001",
3   "_id": "dnX0e4Q89PpvP0Dh1Cxe",
4   "_version": 1,
5   "_score": 0,
6   "_source": { ... },
141 },
142   "_fields": { ... },
345 }
346 }
```

- ¿Qué es lo que se quiere hacer con esta herramienta?
 - Detectar direcciones IP asociadas a actividades ilícitas en la web.
 - En primera instancia solamente se hará con direcciones ip de bots que rastrean la web con otros fines. Por ejemplo, indexado de contenido a los principales motores de búsqueda en la web.
- ¿Cómo hacer consulta sobre los índices creados?
 - Utilizaremos por el momento el parámetro "q" el cual permite realizar consultas con la sintaxis de lucene.
 - Utilizaremos el parámetro query del método search de la clase elasticsearch (<https://elasticsearch-py.readthedocs.io/en/v8.5.1/api.html?highlight=elasticsearch%20search#elasticsearch.Elasticsearch.search>)
 -
- ¿Entiendo la documentación que proporciona elasticsearch client en relación a la utilización de los métodos y específicamente la simbología para la representación de los parámetros de métodos?
- Estos son los campos necesarios para realizar una consulta:
 - Campo que contiene la dirección ip : [hits][_source][nginx][access][remote_ip_list]
 - Campo que contiene el verbo de la petición http:[hits][_source][http][request][method]
 - Campo que contiene el código de respuesta: [hits][_source][http][response][status_code]
- La gran duda que tengo es por qué la cantidad de documentos encontrados siempre es de 10.
 - Solución: el parámetro size del cliente Python para elastic permite modificar dicho comportamiento.