# Ejercicio 1 - Instalación y configuración de ES

|  |  |
| --- | --- |
| Apellidos, Nombre: | Campos de Padua, Alfonso |
| Puntuación Final: |  |

## Pregunta 1 [Puntuación 2]

### Explica de forma detallada si recomendarías hacer uso de ElasticSearch como un sistema de base de datos tradicional ¿Qué ventajas o qué inconvenientes puede tener este uso?

No lo recomiendo. En líneas generales un sistema indexador como ELasticSearch lo veo como una herramienta complementaria al sistema de datos en sí, de forma que se puedan indexar las tablas por las columnas más utilizadas y dado que al final esto implica un hasheo, recuperar la información de forma muy rápida - O(1).

Hay que considerar que un sistema SQL sigue aportando ACID y es extremadamente eficiente (30 años de evolución) para los casos de uso para los que esta pensado. Lo ideal es disponer de los datos de forma independiente, sobre todo si sufren frecuentes modificaciones y no interesa reindexar continuamente. Además, habrá procesos que no requieran de indexación y ataquen de forma directa a la BBDD, por lo que considero mejor que sean sistemas diferentes.

La ventaja seria tenerlo todo auto contenido, pero parece una mala praxis. Habría que evaluar el caso de uso concreto por si la excepción tuviera sentido.

## Pregunta 2 – [Puntuación 2]

### Explica las ventajas o inconvenientes que puede tener configurar los archivos de log y de datos en discos diferentes

La ventaja obvia es que al separar el acceso no se satura el I/O del disco. Lógicamente, esto se puede gestionar de forma muy fácil con 2 puntos de montaje mapeados a LVM sobre soportes físicos diferentes, por lo que quizás la única desventaja es que implica tener más de un punto de montaje mapeado a diferente disco físico para que tenga sentido.

## Pregunta 3- [Puntuación 0,5]

### ¿Cuál es el mínimo número de réplica para soportar la caída de dos servidores en un clúster de 5 nodos?

En el peor escenario, se caen 2 servidores que contienen el dato A. Para seguir teniendo acceso a ese dato, tendría que estar copiado en 3 sitios. Por tanto, el factor de réplica seria 2, teniéndose el dato original y 2 réplicas.

## Pregunta 4- [Puntuación 0,5]

### Describe los pasos a seguir para configurar un nodo de ES para que tenga una memoria máxima de 4GB.

En el fichero de configuración elasticsearch.yml en la sección de memoria debemos activar la variable bootstrap.mlockall: true. Además, se debe setear la variable de entorno ES\_MAX\_MEM a 4 GB.

## Pregunta 5 – [Puntuación 0.5]

### ¿Cómo configurarías ES para reservar durante el arranque los 4GB de memoria reservada?

Ademas de los pasos descritos en el paso anterior, se debe setear la variable de entorno ES\_MIN\_MEM a 4 GB.

## Pregunta 6 – [Puntuación 0,5]

### Mostrar, mediante el uso del API REST de ElasticSearch, que la memoria reservada está siendo utilizada

GET \_cluster/stats?human&pretty

Y se pueden ver tanto el límite de ram para el heap como el uso actual en la sección nodes/jvm/men.

## Pregunta 7 – [Puntuación 1]

### Mostrar, mediante el uso del API REST, el tamaño libre en disco para un nodo de ES

Todos los nodos: GET \_nodes/stats

Un nodo: GET \_nodes/VRTVW1/stats/fs?human&pretty

## Pregunta 8 – [Puntuación 1]

### Configura ES para que escriba información de depuración en los ficheros de log

PUT \_cluster/settings {"transient":{"logger.\_root":"DEBUG"}}

Adicionalmente, podemos editar los ficheros logging.yml de cada nodo y cambiar el nivel de traza de la variable es.logger.level: INFO a DEBUG.

## Pregunta 9 – [Puntuación 1]

### Configura ES para que escriba cualquier consulta ejecutada contra el nodo en los ficheros de log

Parece que la única forma de logar las querys es abusar la funcionalidad provista por el módulo de slowlog.

Dinámicamente se puede hacer para cada índice, mediante:

PUT /\_cluster/settings

{

"transient": {

"index.search.slowlog.threshold.query.warn": "0s",

"index.search.slowlog.threshold.query.info": "0s",

"index.search.slowlog.threshold.query.debug": "0s",

"index.search.slowlog.threshold.query.trace": "0s"

}

}

Además se puede modificar para cada nodo editando el fichero de logging.yml y añadiendo:

index.search.slowlog.threshold.query.warn: 0s

index.search.slowlog.threshold.query.info: 0s

index.search.slowlog.threshold.query.debug: 0s

index.search.slowlog.threshold.query.trace: 0s

También es necesario cambiar el valor de:

index.search.slowlog: false a true.

## Pregunta 10 – [Puntuación 1]

### ¿Qué hace ElasticSearch si en un clúster de 3 nodos de datos habilitamos un factor de replicación de 3?

En este caso las réplicas no se asignaran y el estado del Cluster permanecerá amarillo. El número de réplicas debe ser n – 1 del número de nodos como máximo.