**Báo cáo Lưu Trữ Và Xử Lý Dữ Liệu Lớn**

**Lab3:MongoDB, Elasticsearch**

**Nhóm**:LoveData

**Thành viên**:

Phương Trung Kiên 20183776

Lê Nguyễn Thành Chung 20183696

Đào Trọng Hiếu 20183739

HIM PHON 20170283

1. **Yêu cầu:**

-Cài đặt NoSQL: Mongodb hoặc Elasticsearch. Yêu cầu: có tối thiểu 3 máy vật lý (máy ảo) tạo thành cluster, chụp hình minh chứng, đưa dữ liệu 10M~1G lên NoSQL và đưa minh chứng được lưu phân tán ra sao.

1. **Mô tả cài đặt:**

-Sử dụng 3 máy ảo từ 2 bài thực hành trước với các địa chỉ IP:

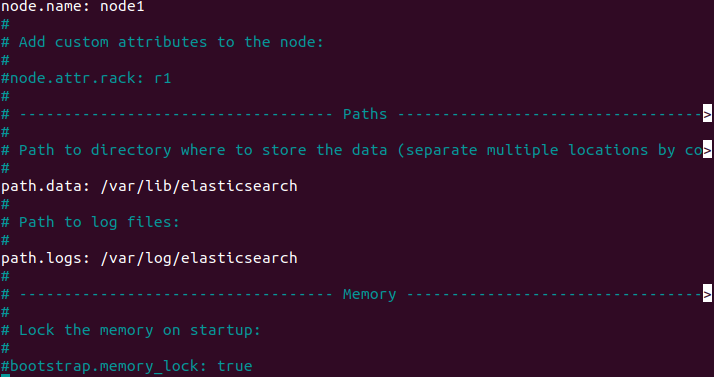
* 192.168.3.11:nodems
* 192.168.3.12:node1
* 192.168.3.13:node2

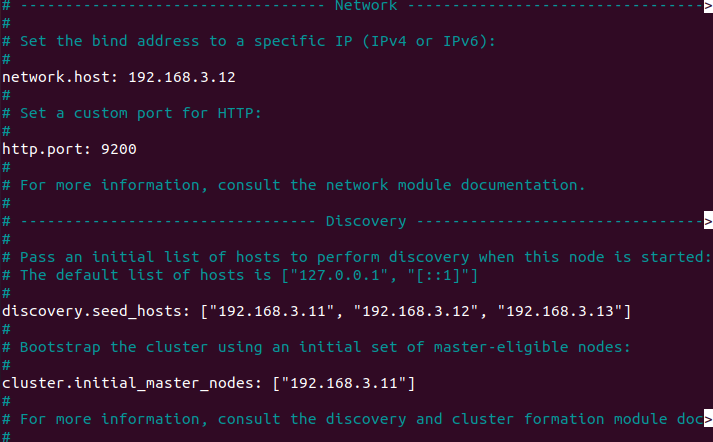
-Cài đặt Elasticsearch trên các node

-Cấu hình file elasticsearch.yml trên các node :

+Cấu hình các thông số trên các máy:

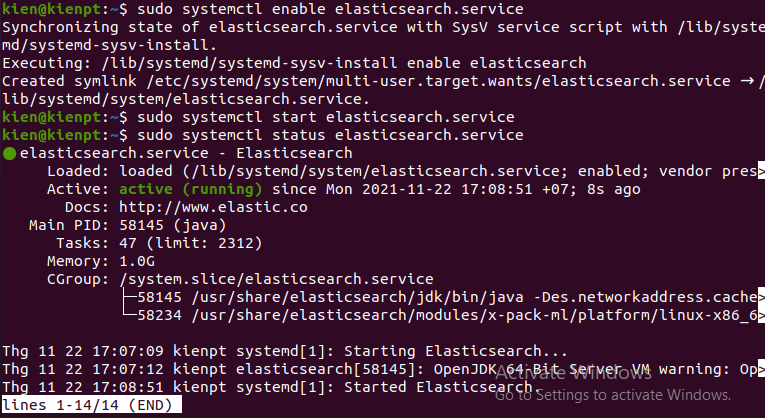
* node.name.
* network.host.
* discovery.seed\_hosts.
* cluster.initial\_master\_nodes

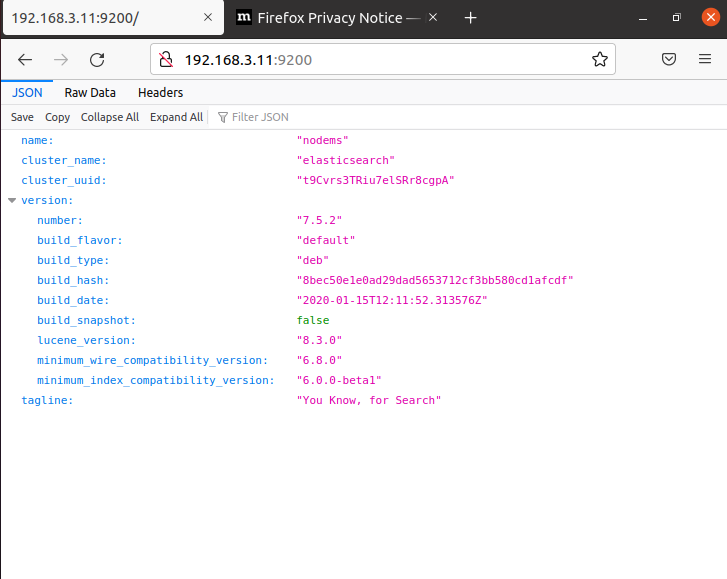




-Khởi chạy elasticsearch:

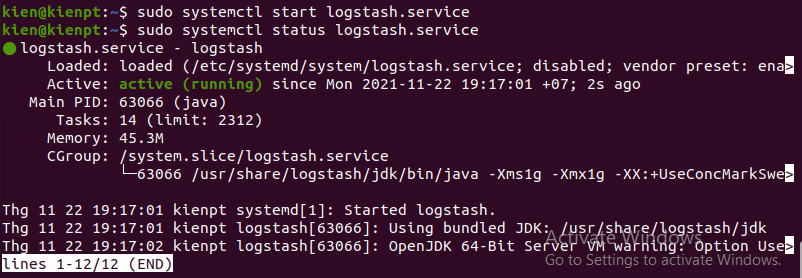
* **Sudo systemctl enable elasticsearch.service**
* **Sudo systemctl start elasticsearch.service**





-Cài đặt logstart để lưu trữ dữ liệu:

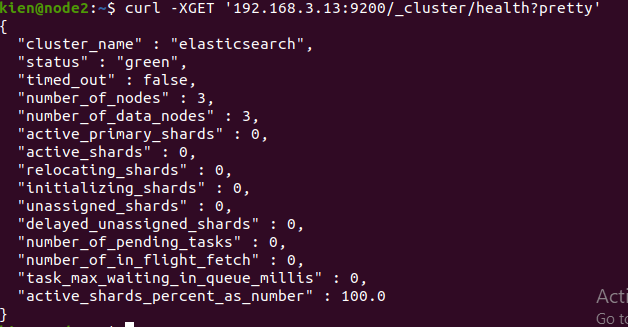
* **wget -qO - https://artifacts.elastic.co/GPG-KEY-elasticsearch | sudo apt-key add -**
* **echo "deb https://artifacts.elastic.co/packages/7.x/apt stable main" | sudo tee -a /etc/apt/sources.list.d/elastic-7.x.list**
* **sudo apt update**
* **sudo apt-get install logstart**

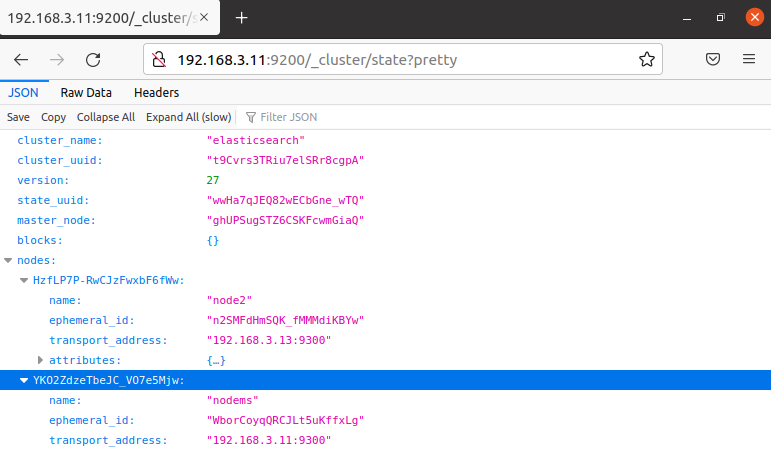


-Tạo 1 file trong có tên data3.conf trong /etc/logstash/conf.d:



1. **Kết quả:**

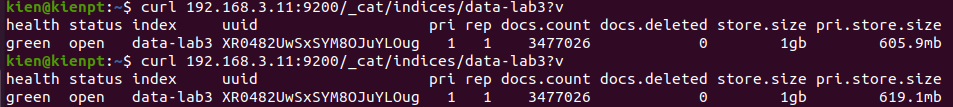
-Kết quả khi cài đặt ở 3 máy:  




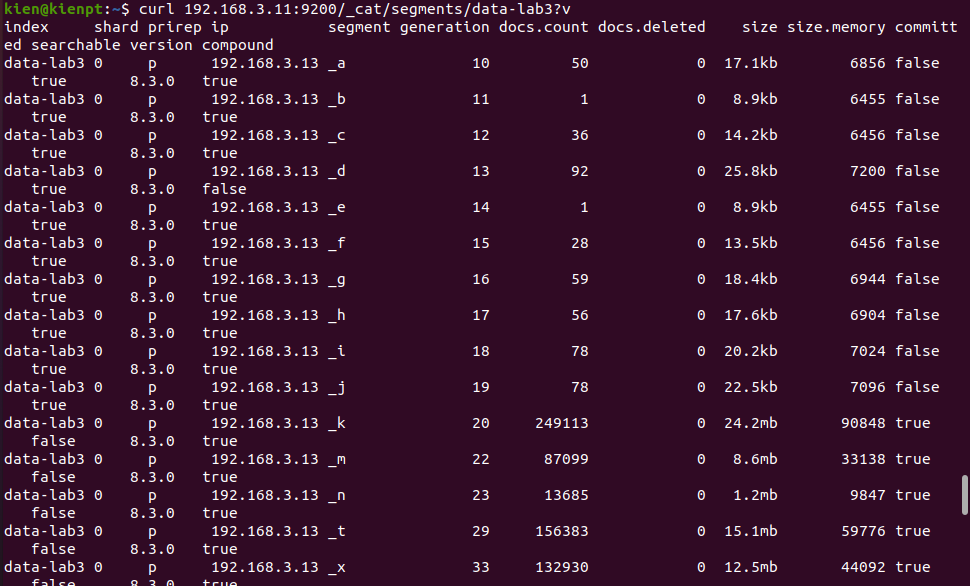
-Kết quả sau khi chạy file data3.conf để lưu trữ dữ liệu:

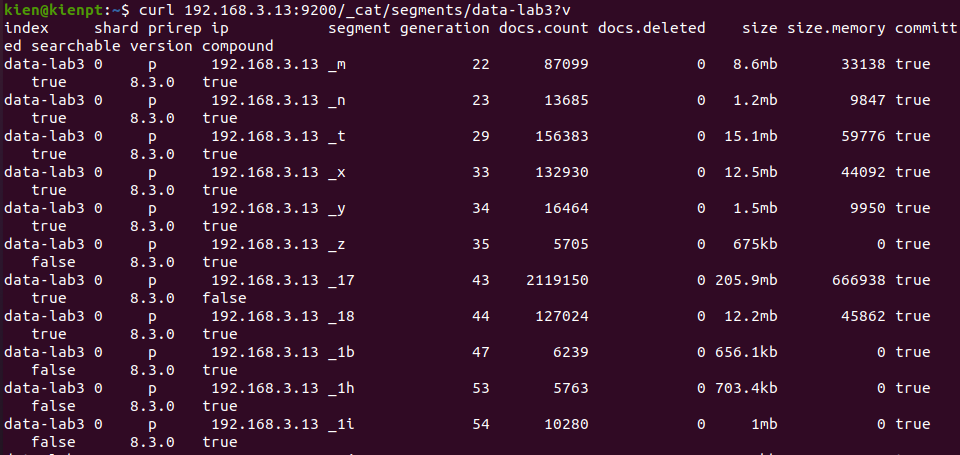
-**/usr/share/logstash/bin/logstash -f /etc/logstash/conf.d/crimes.conf**

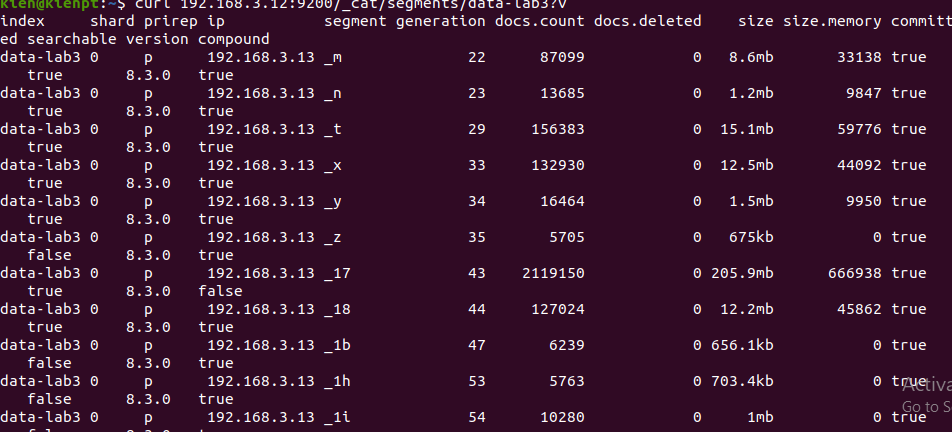
**-crul 192.168.3.11:9200/\_cat/indices/data-lab3?v**

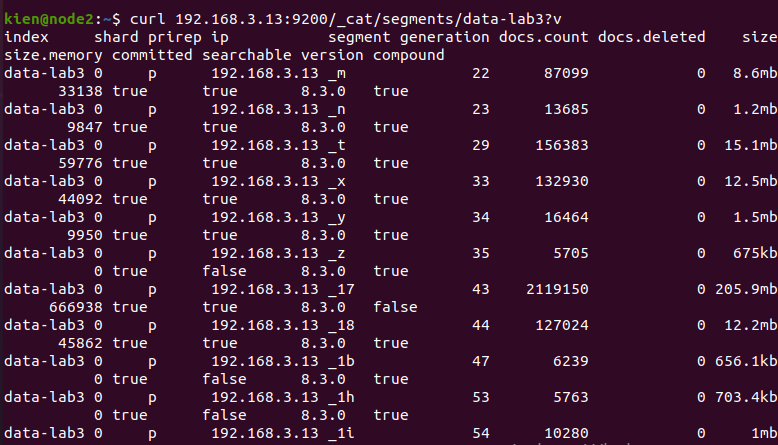


-Dữ liệu phân tán:









-Kết quả hiển thị trên web:

