

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI  
VIỆN CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG



BÁO CÁO MÔN HỌC  
**Lưu trữ và xử lý dữ liệu lớn**

LAB03: Install Elasticsearch

Nhóm: **RHUST**

Phan Thành Đạt	20173001
Hoàng Văn Chương	20172984
Đỗ Minh Vũ	20173471
Nguyễn Thị Bắc	20172963

## 1) Cài đặt cụm Elasticsearch

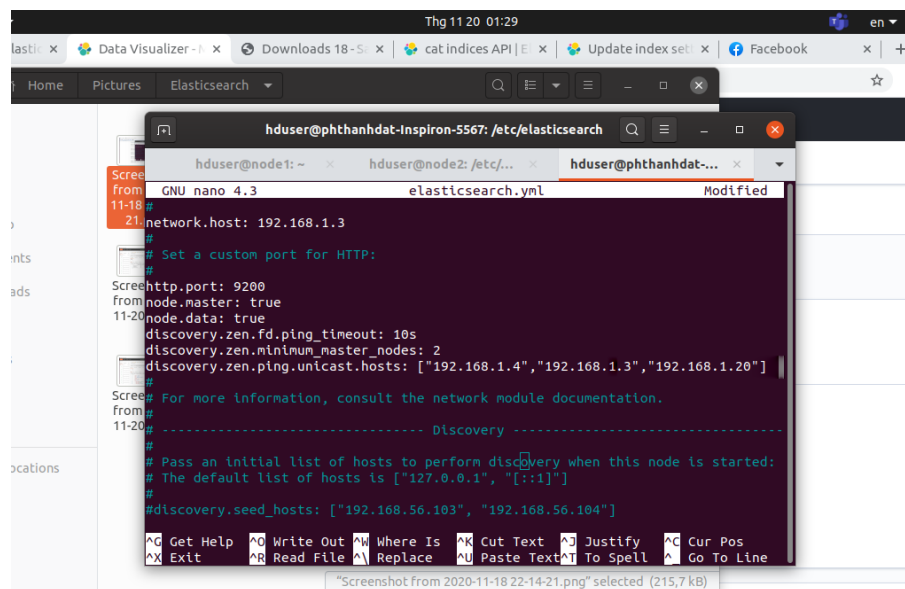
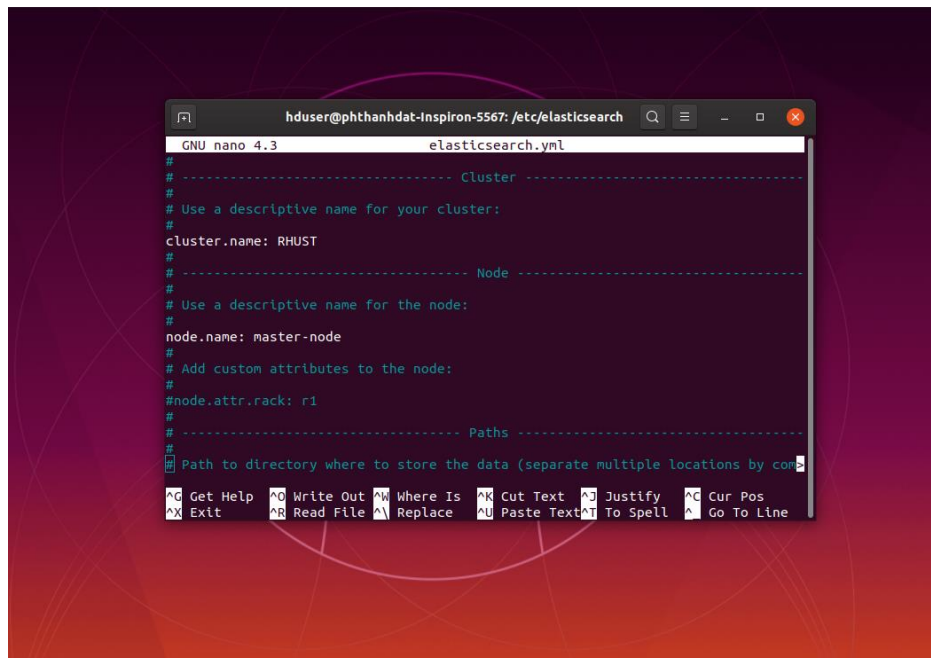
3 máy vật lý:

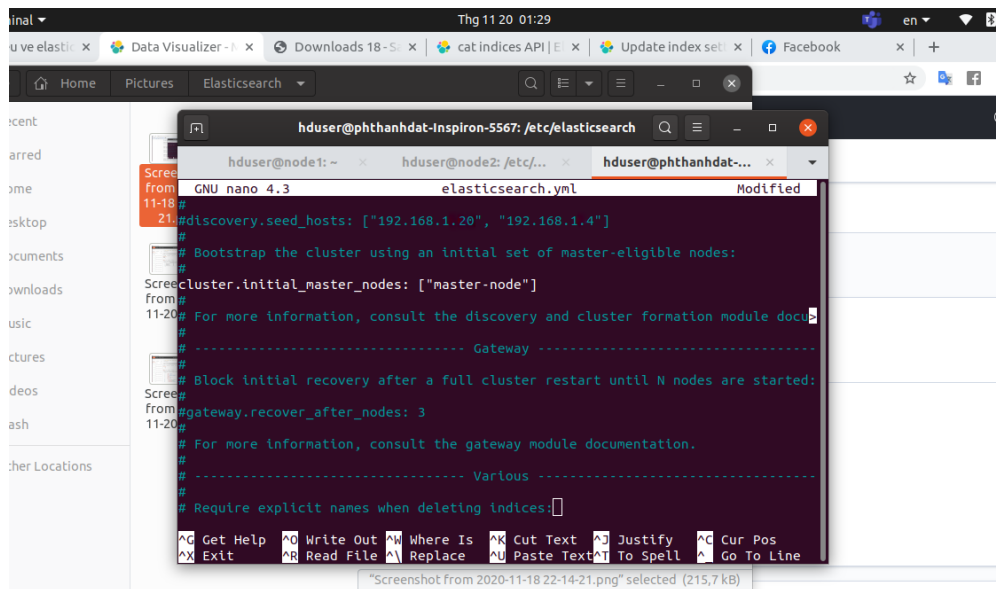
192.168.1.3 : master-node

192.168.1.20:node1

192.168.1.4:node2

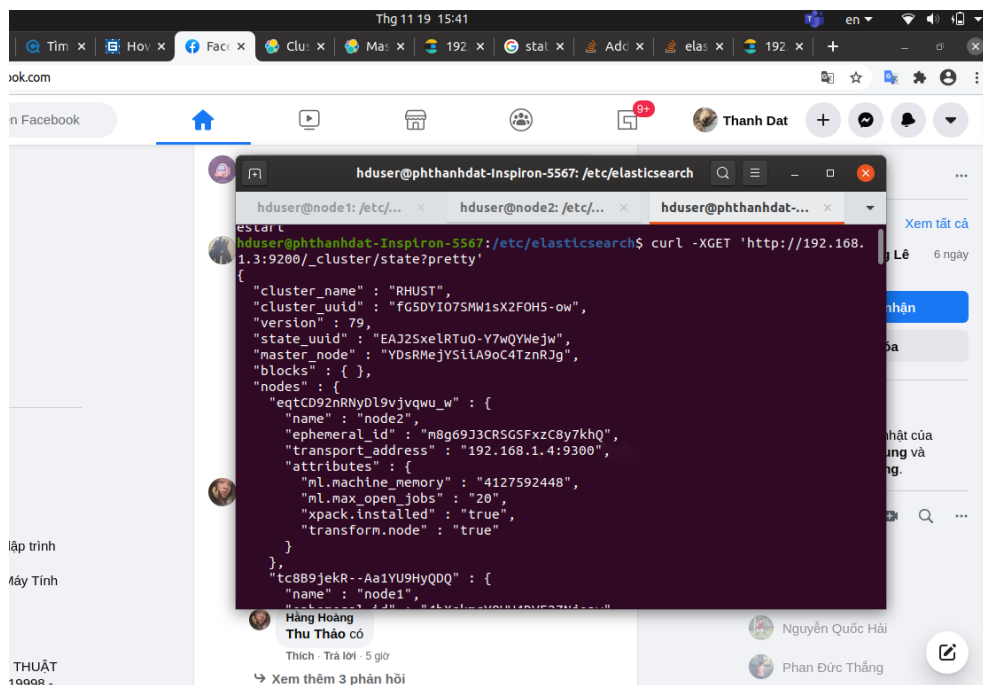
Tên cụm: RHUST



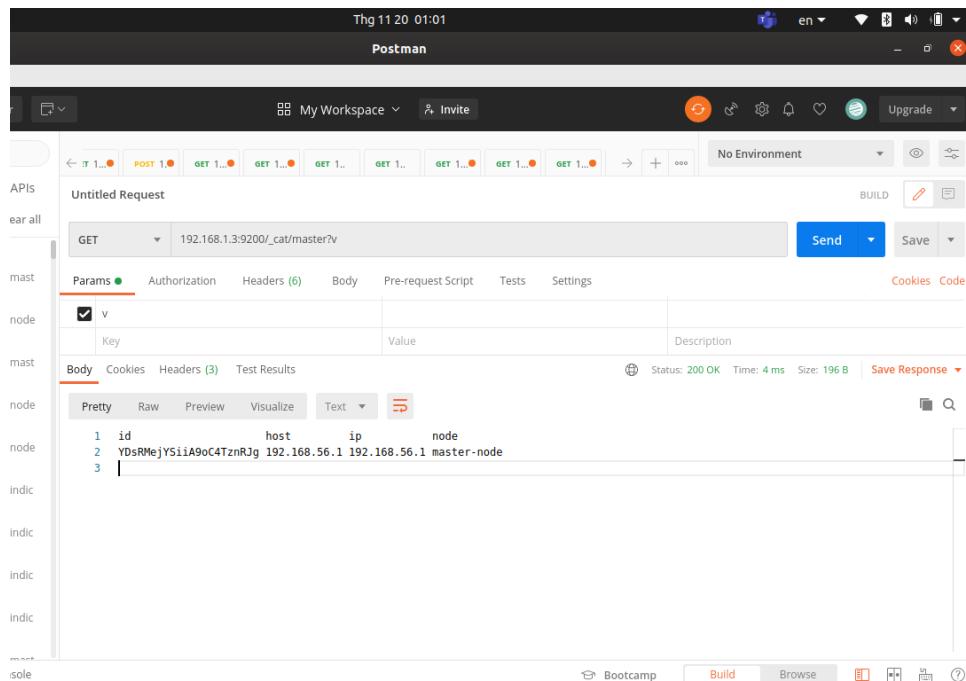


Chạy lần lượt các máy bằng command line: **sudo service elasticsearch start**

Xem thông tin cụm: **curl -XGET 'http://192.168.1.3:9200/\_cluster/state?pretty'** (method GET)





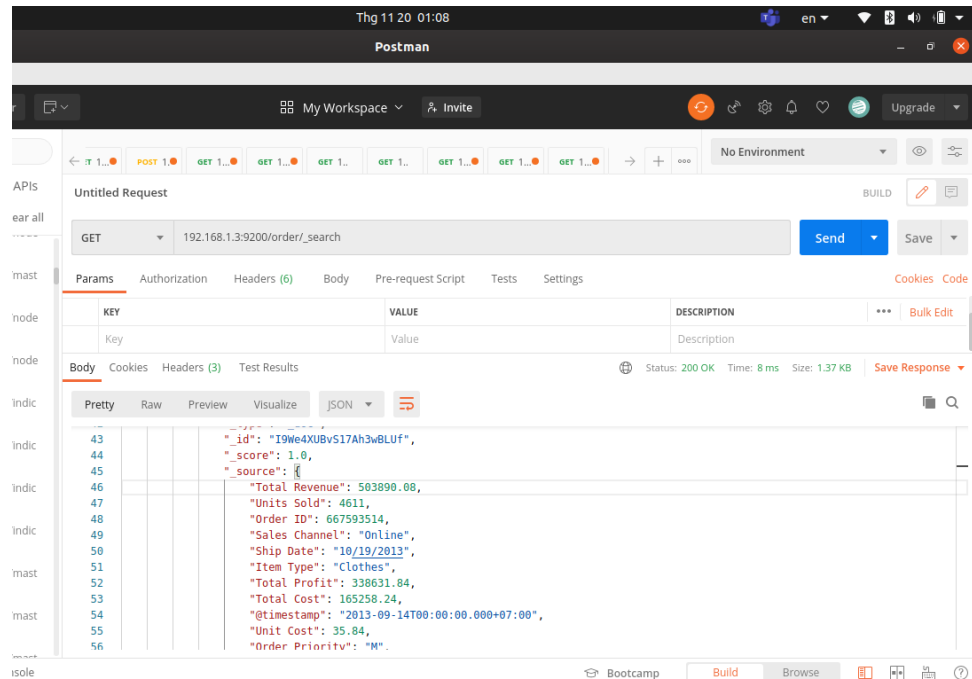


2) API thêm, sửa, xóa, lấy thông tin.

a) Method GET:

Lấy toàn bộ document của 1 index:

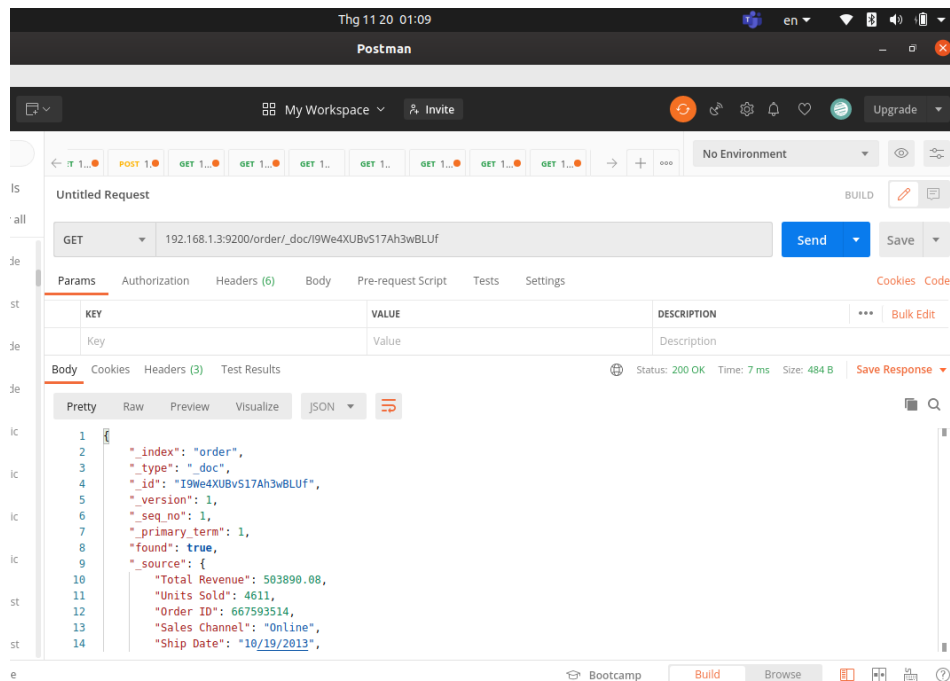
192.168.1.3:9200/\_index/\_search



Lấy thông tin 1 document:

GET /index\_name/type\_name/id

192.168.1.3:9200/index\_name/type\_name/id



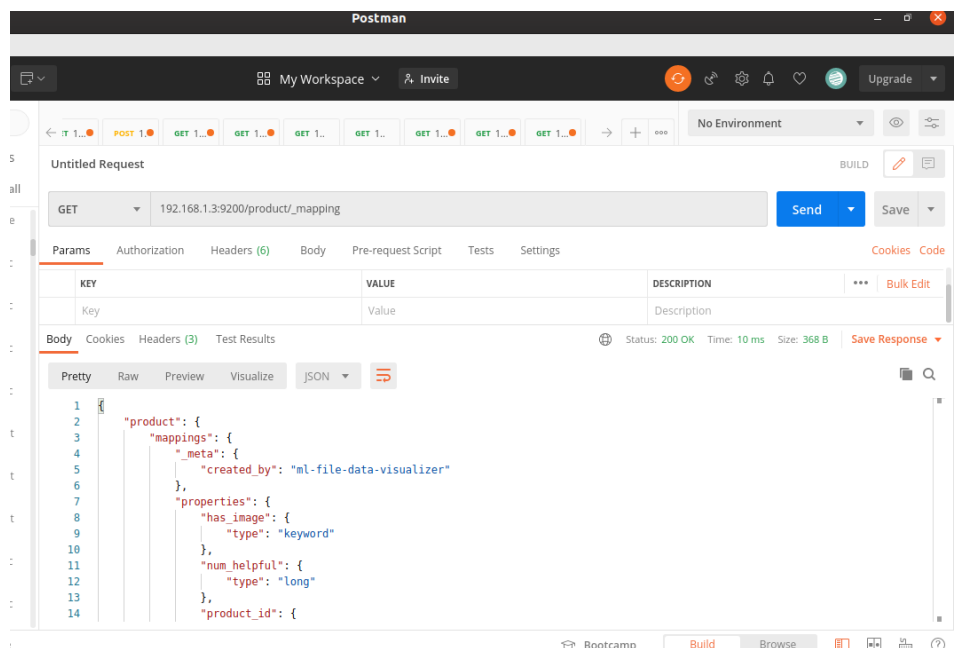
Tìm kiếm tất cả các document trong 1 type:

GET /index\_name/type\_name/\_search

Tìm kiếm dữ liệu theo tham số đầu vào:

GET /\_search?q=tham\_so

Ngoài ra, có thể xem thông tin chi tiết về index:



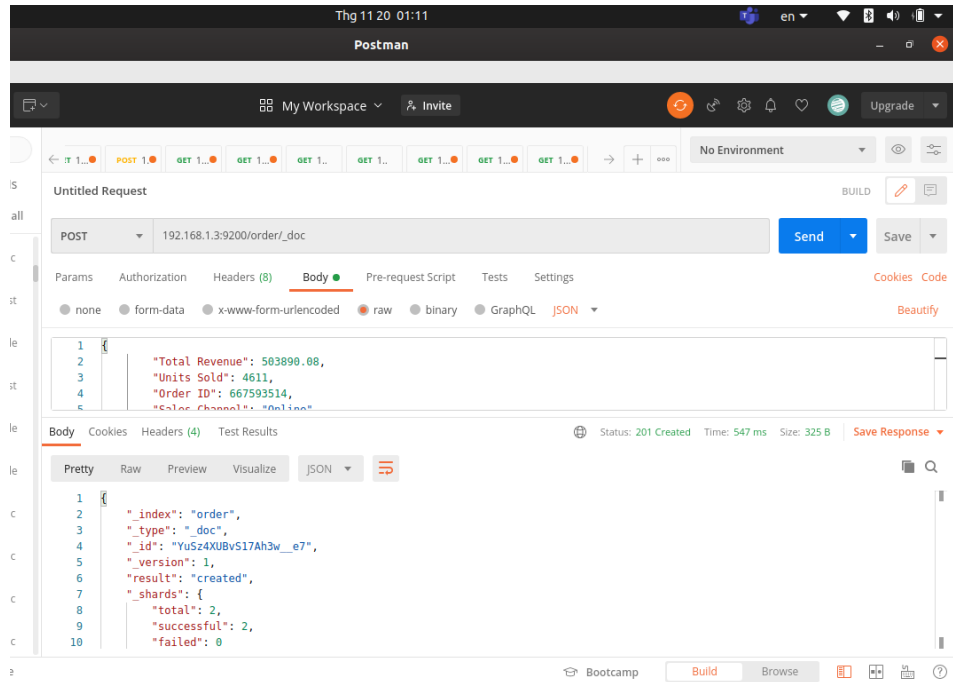
b) Method POST:

POST /index\_name/type\_name/id

```

{
  "field" : "value"
}
  
```

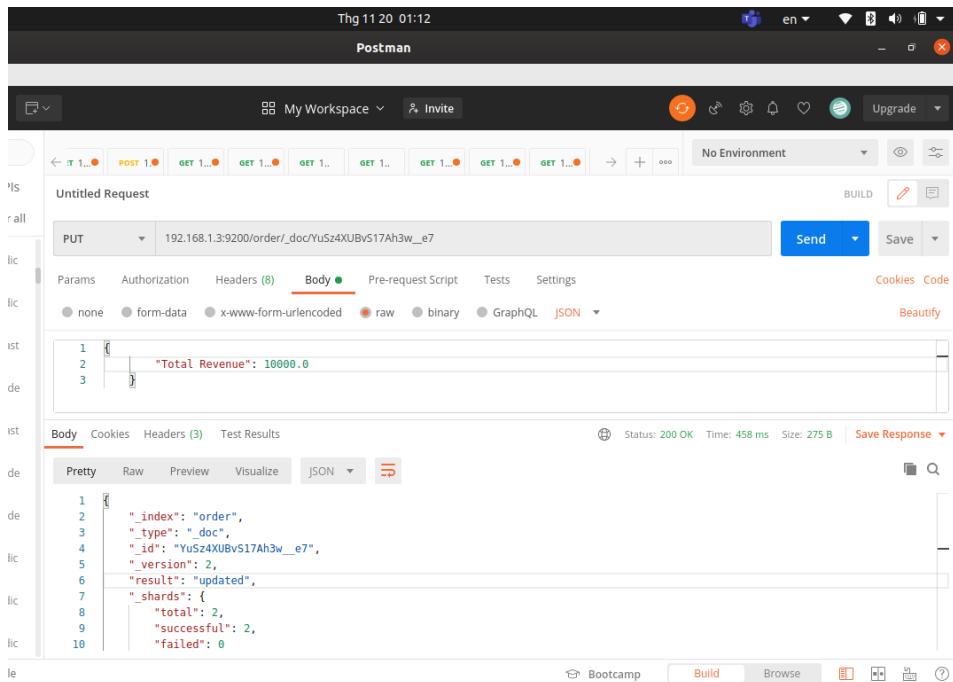
}



c) Method PUT:

PUT /index\_name/type\_name/id

```
{
  "field": "data"
}
```

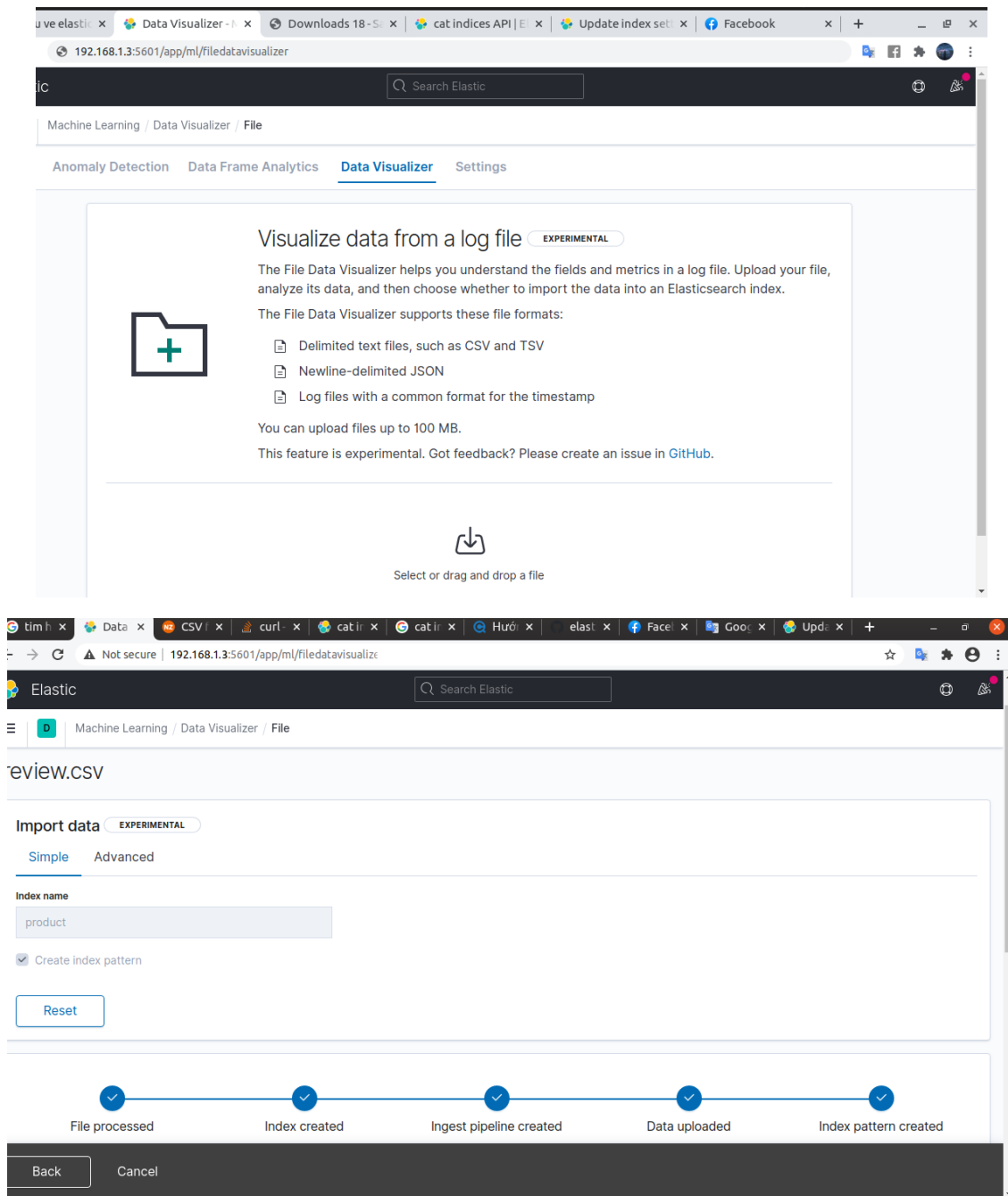


d) Method DELETE

DELETE /index\_name/type\_name/id

3) Xem thông tin về các index ~ 1GB data

Cài Kibana để up file csv hoặc json kích thước lớn:



Lấy thông tin các index trong cluster:



Postman interface showing a GET request to `192.168.1.3:9200/_cat/indices?v`. The response status is 200 OK. The response body is a table with 10 rows of index information.

	health	status	index	uuid	pri	rep	docs.count	docs.deleted	store.size	pri.store.size
1	green	open	product	0hXT4jwrQcQLifb79Wwz1A	1	1	12223	0	12mb	6.2mb
2	green	open	5000_sales	sn5L-o3tRfi8_T7b3nvuGg	1	1	500000	0	203.7mb	101.7mb
3	green	open	bigdata	L6iD4m28QNC2UWVF-hIAEW	1	1	1	0	7.7kb	3.8kb
4	green	open	.apm-agent-configuration	8596jXPrLiassFIUN8DWA	1	1	0	0	416b	208b
5	green	open	order_5	7i23gvKWReH-Q3YtNZnPw	1	1	500000	0	200.1mb	101.8mb
6	green	open	order_1	Vp2aIAmSXGSnJ_0mGirvw	1	1	500000	0	203.6mb	101.8mb
7	green	open	.kibana_1	dTpooHnzSuulDE0pT0seng	1	1	30	26	20.8mb	10.4mb
8	green	open	order_2	cebn4z4vTPqxUz3wswNjrw	1	1	500000	0	205.2mb	101.8mb
9	green	open	order_3	8yTWBdepRZ0ydA-40u9M6Q	1	1	500000	0	204.5mb	101.7mb

Postman interface showing a GET request to `192.168.1.3:9200/_cat/indices?v`. The response status is 200 OK. The response body is a table with 16 rows of index information.

	health	status	index	uuid	pri	rep	docs.count	docs.deleted	store.size	pri.store.size
7	green	open	order_1	Vp2aIAmSXGSnJ_0mGirvw	1	1	500000	0	203.6mb	101.8mb
8	green	open	.kibana_1	dTpooHnzSuulDE0pT0seng	1	1	30	26	20.8mb	10.4mb
9	green	open	order_2	cebn4z4vTPqxUz3wswNjrw	1	1	500000	0	205.2mb	101.8mb
10	green	open	order_3	8yTWBdepRZ0ydA-40u9M6Q	1	1	500000	0	204.5mb	101.7mb
11	green	open	order_4	lJo6hzhfTEW1SDLua8eH0A	1	1	500000	0	201mb	103mb
12	green	open	testdata	bDy_N63tQW05bXlXU8pdFA	1	1	2884	0	4.8mb	2.3mb
13	green	open	.apm-custom-link	oUNuk0PDQJ-90EFkSpzgrw	1	1	0	0	416b	208b
14	green	open	.kibana_task_manager_1	zJtm2hZKT75TYzVzDLmrNg	1	1	5	190	277.6kb	115.8kb
15	green	open	.kibana-event-log-7.10.0-000001	myX-EQjtT4MTsa8DlB5Hlg	1	1	1	0	11.2kb	5.6kb

Health: tình trạng của index (green,yellow,red)

Status : trạng thái

Index : tên index

Pri: số lượng phân mảnh

Rep: số lượng bản sao

Doc.count: số lượng document có trong index

Store.size: Kích thước của index tính cả các bản sao = pri.store.size \* số lượng bản sao

Pri.store.size: kích thước của 1 bản chính.

Thay đổi số lượng bản sao: method PUT, truyền tham số number\_of\_replicas ở body

Thg 11 20 01:37

Postman

Runner

My Workspace

Invite

Upgrade

PUT

192.168.1.3:9200/product/\_settings

Send

Save

Params

Authorization

Headers (8)

Body

Pre-request Script

Tests

Settings

none

form-data

x-www-form-urlencoded

raw

binary

GraphQL

JSON

1

2

3

4

5

{

"index": {

"number\_of\_replicas": 3

}

}

Body

Cookies

Headers (3)

Test Results

Status: 200 OK

Time: 6.21 s

Size: 157 B

Save Response

Pretty

Raw

Preview

Visualize

JSON

1

2

3

{

"acknowledged": true

}

Console

Build

Browse

Postman

Runner My Workspace Invite No Environment

Untitled Request

PUT 192.168.1.3:9200/\_sales/\_settings Send Save

Params Authorization Headers (8) Body Pre-request Script Tests Settings Cookies Code

none form-data x-www-form-urlencoded raw binary GraphQL JSON Beautify

```
1 {
2   "index":
3   "number_of_replicas":2
4 }
5
```

Body Cookies Headers (3) Test Results Status: 200 OK Time: 4.55 s Size: 157 B Save Response

Pretty Raw Preview Visualize JSON

```
1 {
2   "acknowledged": true
3 }
```

Console Bootcamp Build Browse

Postman

Runner My Workspace Invite No Environment

Untitled Request

GET 192.168.1.3:9200/\_cat/indices?v Send Save

Params Authorization Headers (6) Body Pre-request Script Tests Settings Cookies Code

Query Params

KEY	VALUE	DESCRIPTION	*** Bulk Edit
<input checked="" type="checkbox"/> v			
Key	Value	Description	

Body Cookies Headers (3) Test Results Status: 200 OK Time: 66 ms Size: 822 B Save Response

Pretty Raw Preview Visualize Text

```
1 health status index uuid pri rep docs.count docs.deleted store.size pri.store.size
2 yellow open product 0hXT4jwr0CqLifb79Wwz1A 1 3 12223 0 12mb 6.2mb
3 yellow open 5000 sales sn5L-o3trFi0-T7b3nvuGg 1 2 5000000 0 203.7mb 101.7mb
4 green open bigdata L6id4m28QNC2UwVf-hIAEw 1 1 1 0 7.7kb 3.8kb
5 green open .apm-agent-configuration 8596jXPpRLiazsFIUN8DWA 1 1 0 0 416b 200b
6 green open order_5 7iz3gvKwRweH-03YtNZnPw 1 1 500000 0 204.2mb 101.8mb
7 green open order_1 Vp2a1AmSSXG5nJ-0mGlrw 1 1 500000 0 203.6mb 101.8mb
8 green open .kibana 1 dTpoohnz5uu1dE0t0senq 1 1 31 0 41.6mb 20.8mb
```

Console Bootcamp Build Browse