TRIỂN KHAI THỬ NGHIỆM HỆ THỐNG QUẢN LÝ VÀ PHÂN TÍCH LOG TẬP TRUNG ELK STACK

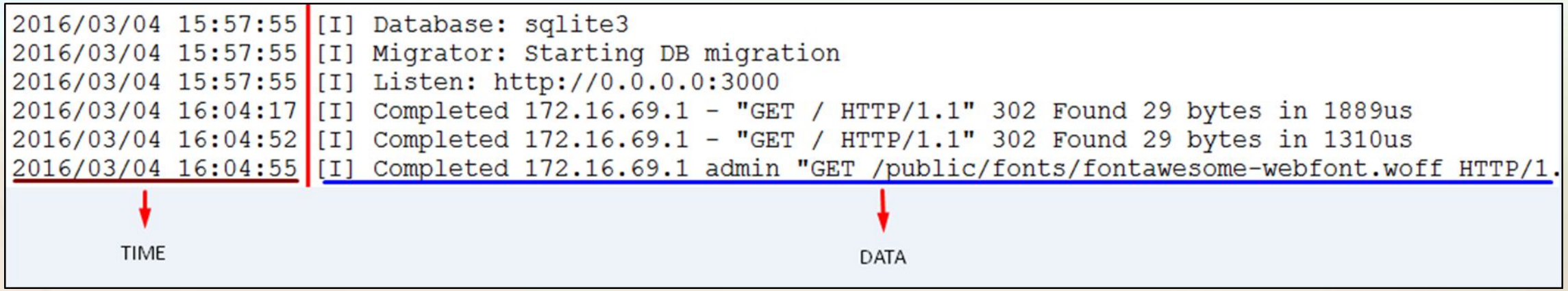
Trình bày: Nguyễn Trọng Tấn

NỘI DUNG

1. Tìm hiểu về LOG
2. Tìm hiểu về ELK stack
3. Triển khai thử nghiệm
4. Thảo luận

1.1 LOG là gì?

- LOG là thông tin về hoạt động của hệ thống hoặc các dịch vụ.
- LOG có định dạng chung là: thời gian + dữ liệu.



1.2 Tại sao cần lưu trữ và phân tích LOG?

- Giúp theo dõi tình trạng hoạt động của hệ thống hoặc các dịch vụ
- Phân tích LOG để tìm nguyên nhân khi có sự cố. Giúp khắc phục sự cố nhanh hơn.
- Phát hiện và dự đoán các vấn đề có thể xảy ra cho hệ thống.

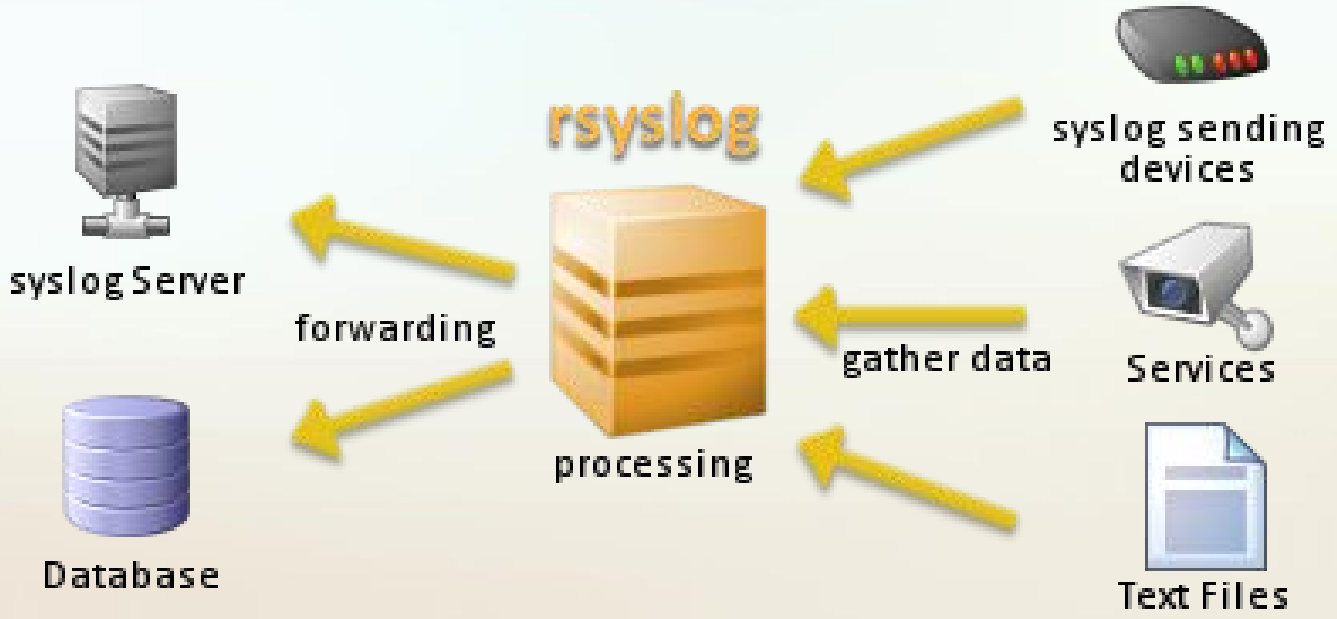
1.3 Các giải pháp lưu trữ LOG

- Lưu trữ phân tán: Lưu LOG tại local, vị trí mặc định là thư mục /var/log

Log file	Description
/var/log/boot.log	Boot log information.
/var/log/httpd	Apache web server log.
/var/log/messages	Post boot kernel information.
/var/log/auth.log	User authentication log.
/var/log/dmesg	System boot up messages.
/var/log/mail.log	Mail server log.
/var/log/Xorg.0.log	X Server log.

1.4 Các giải pháp lưu trữ LOG

- Lưu trữ tập trung



Source:https://linux.cn/article-4835-1.html

1.5 Hạn chế trong lưu trữ LOG mặc định

- Tra cứu và phân tích LOG phải sử dụng lệnh: cat, grep, more, head, tail, awk,…
- Chưa có tính năng tự động phân tích LOG.
- Chưa chuyển các dữ liệu thô thành các bản tin dễ đọc, dễ hiểu.

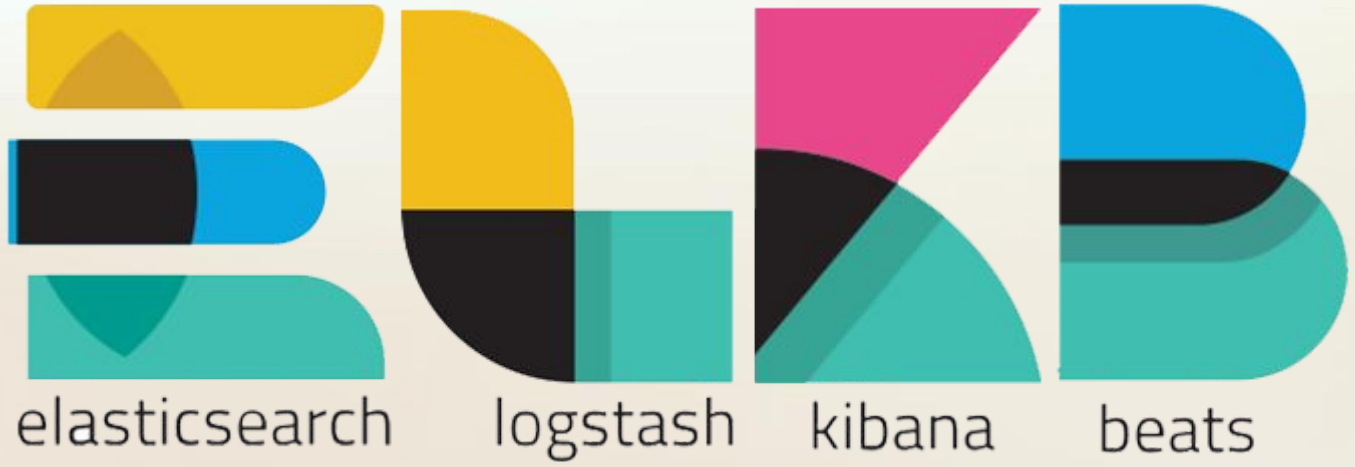
➔ ELK stack là một công cụ để giải quyết các hạn chế trên.

NỘI DUNG

1. Tìm hiểu về LOG
2. Tìm hiểu về ELK stack
3. Triển khai thử nghiệm
4. Thảo luận

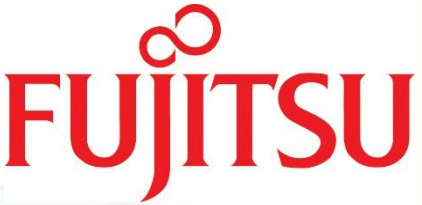
2.1 ELK stack là gì?

- ELK stack được tạo lên từ 03 thành phần mã nguồn mở Elasticsearch, Logstash, Kibana; có chức năng thu thập, phân tích, lưu trữ, tìm kiếm, hiển thị dữ liệu.
- Ngoài ra, Beats là thành phần rất quan trọng đi kèm với ELK stack. Beats được cài đặt trên Client với nhiệm vụ gửi dữ liệu về Logstash



Source:https://goo.gl/TKBOqQ

2.2 Ai đang sử dụng ELK stack?

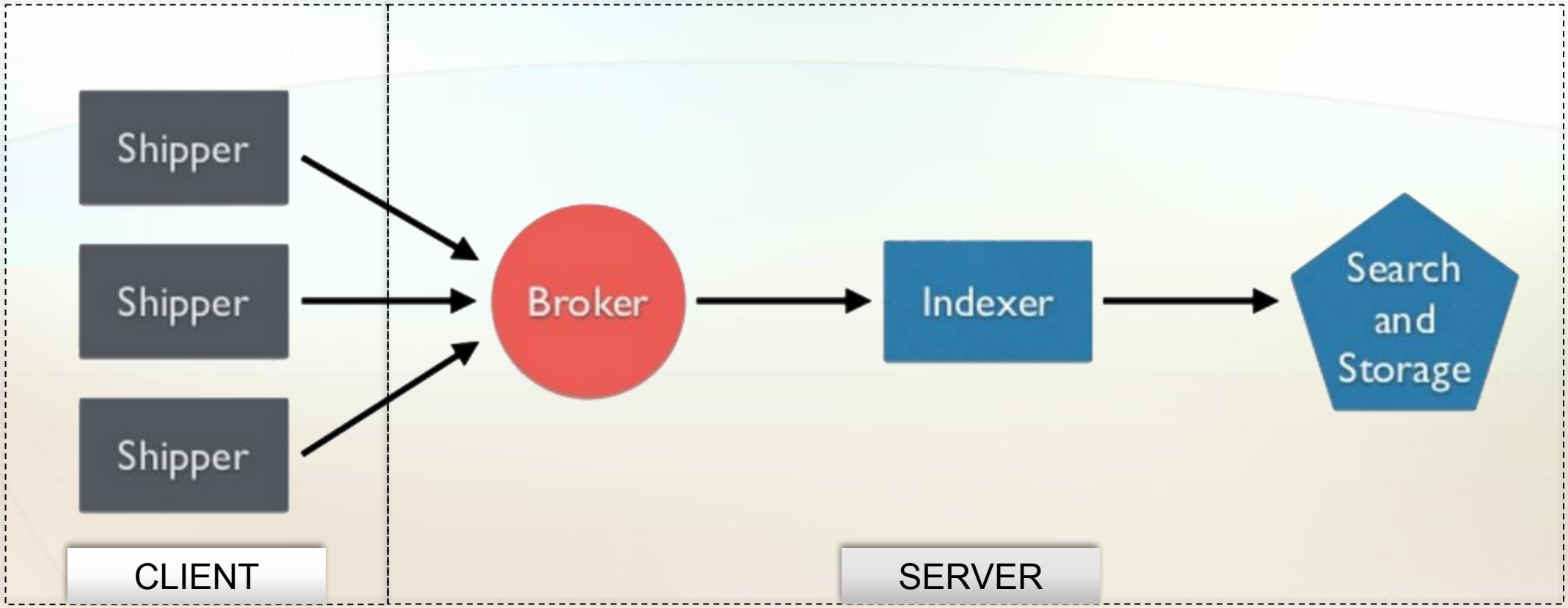


Source:https://www.elastic.co/use-cases

2.3 Đặc điểm của ELK stack

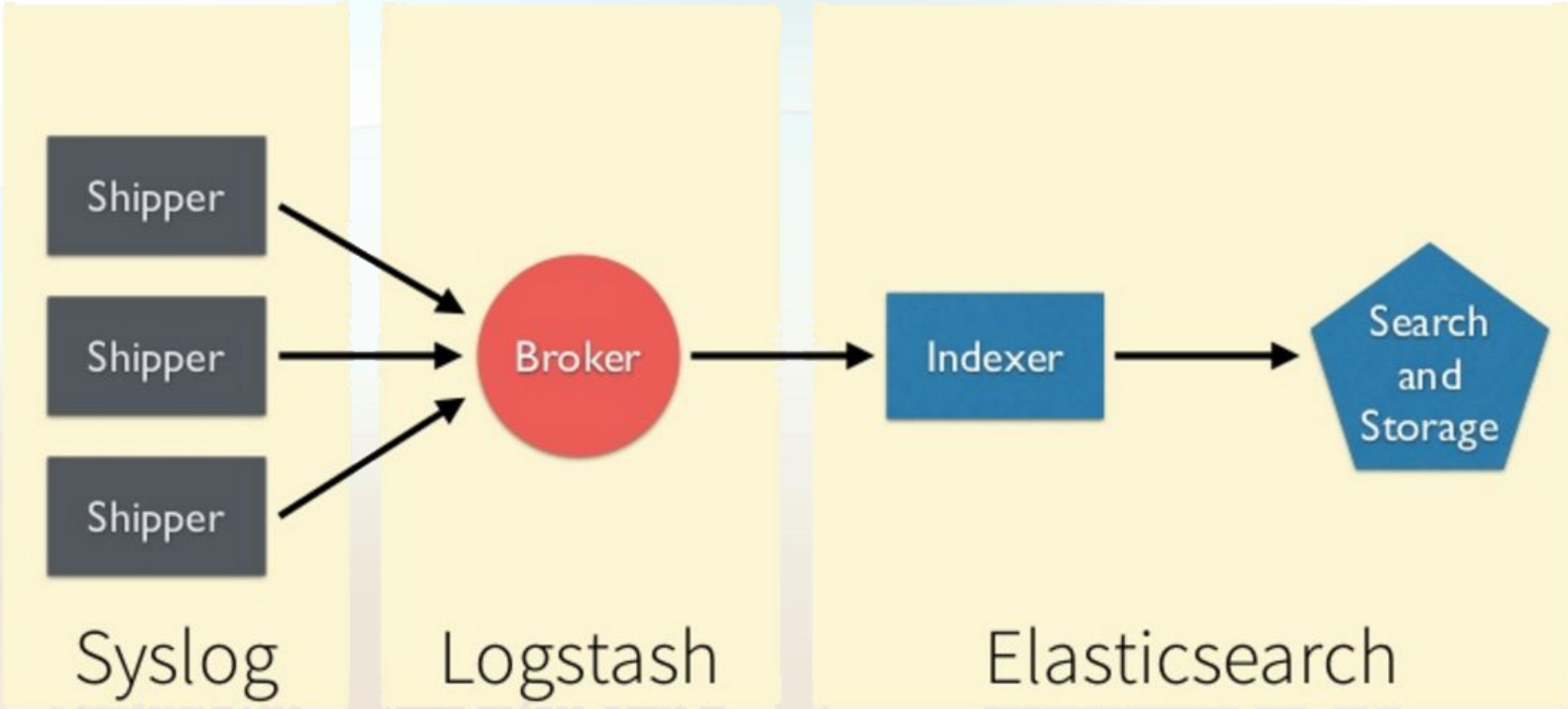
- Lưu trữ, hiển thị, tìm kiếm dữ liệu theo thời gian thực.
- Thu thập, phân tích khối lượng dữ liệu lớn tại một thời điểm.
- Cài đặt, cấu hình, mở rộng và vận hành đơn giản.
- Giao diện tương tác người dùng thân thiện, trực quan.

2.4 Kiến trúc tổng quan



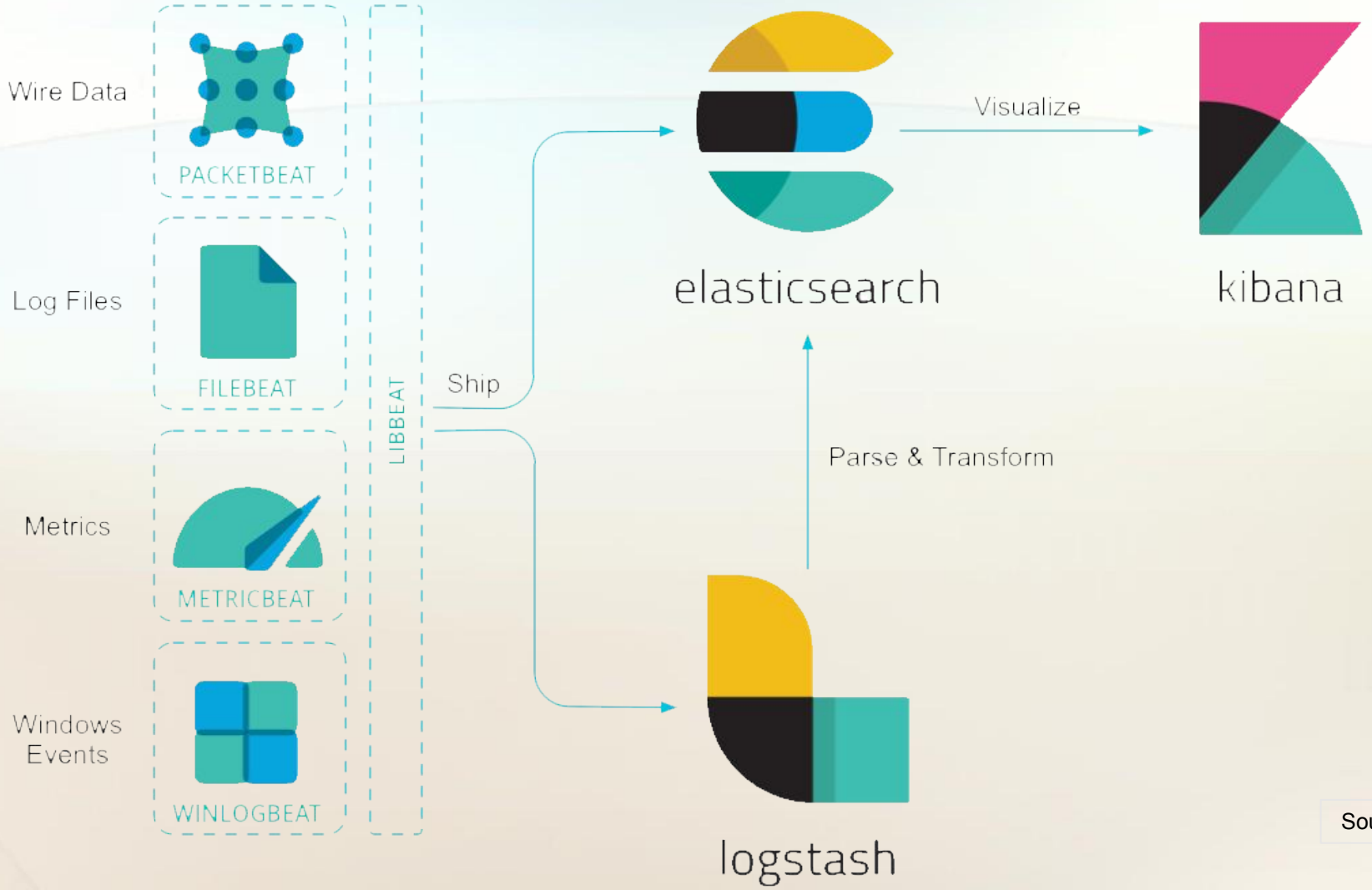
Source:https://goo.gl/zfdHTQ

2.5 Kiến trúc phân lớp



Source:https://goo.gl/zfdHTQ

2.6 Kiến trúc cụ thể



Source:https://goo.gl/h1dCnQ

2.7 Các thành phần của ELK stack

Ship:

Beats – Thu thập dữ liệu trên client và gửi về ELK Server.

Stream:

Logstash – Tiếp nhận, phân tích, chuyển tiếp dữ liệu.

Store:

Elasticsearch – Lưu trữ, đánh chỉ mục, tìm kiếm phân tán thời gian thực.

View:

Kibana – Giao diện người dùng, hiển thị biểu đồ, phân tích cú pháp tìm kiếm.

2.8 Beats

- Beats được viết bằng Golang.
- Beats thu thập dữ liệu trên Client và gửi về ELK Server.
- Dữ liệu được beats thu thập là: file, packet, resource,…
- Mỗi loại beats thực hiện thu thập các loại dữ liệu riêng.
- Dựa vào thư viện libbeat có thể phát triển beats theo yêu cầu.

2.9 Cấu hình Beats

```
filebeat:
  prospectors:
    -
      paths:
        - /var/log/neutron/*.log
      input_type: log
      document_type: syslog
  registry_file: /var/lib/filebeat/registry
output:
  logstash:
    hosts: ["172.16.69.90:5044"]
    tls:
      certificate_authorities: ["/etc/pki/tls/certs/logstash-forwarder.crt"]
shipper:
  name: client1
  tags: [“syslog"]
```

2.10 Logstash

- Xuất nhập nhiều loại dữ liệu
- Tiếp nhận dự liệu
- Phân tích dữ liệu
- Xuất dữ liệu

2.11 Logstash

- **Input**: file, syslog, redis,…
- **Codecs**: plain,…
- **Filters**: grok, geoip,…
- **Output**: elasticsearch, file,…

2.12 Cấu hình Logstash

```
input {
  beats {
    port => 5044
    ssl => true
    ssl_certificate => "/etc/pki/tls/certs/logstash-forwarder.crt"
    ssl_key => "/etc/pki/tls/private/logstash-forwarder.key"
  }
}
```

2.13 Cấu hình Logstash

```
filter {
  if [type] == "syslog" {
    grok { match => {"message" => "%{TIMESTAMP_ISO8601:timestamp}
%{LOGLEVEL:log-level} \[%{DATA:class}\]:%{GREEDYDATA:message}" }}
   }
}
```

2.14 Cấu hình Logstash

```
output {
  elasticsearch {
    hosts => ["localhost:9200"]
    sniffing => true
    manage_template => false
    index => "%{[@metadata][beat]}-%{+YYYY.MM.dd}"
    document_type => "%{[@metadata][type]}"
  }
}
```

2.15 Logstash xử lý dữ liệu

- **Input**:

2016-07-11T23:56:42.000+00:00 INFO [MySecretApp.com.Transaction.Manager]:Starting transaction for session -464410bf-37bf-475a-afc0-498e0199f008

- **Filter**:

grok {match => { "message" => "%{TIMESTAMP_ISO8601:timestamp} %{LOGLEVEL:log-level} \[%{DATA:class}\]:%{GREEDYDATA:message}" }}

- **Output**:

{ "message" => "Starting transaction for session -464410bf-37bf-475a-afc0-498e0199f008",

"timestamp" => "2016-07-11T23:56:42.000+00:00",

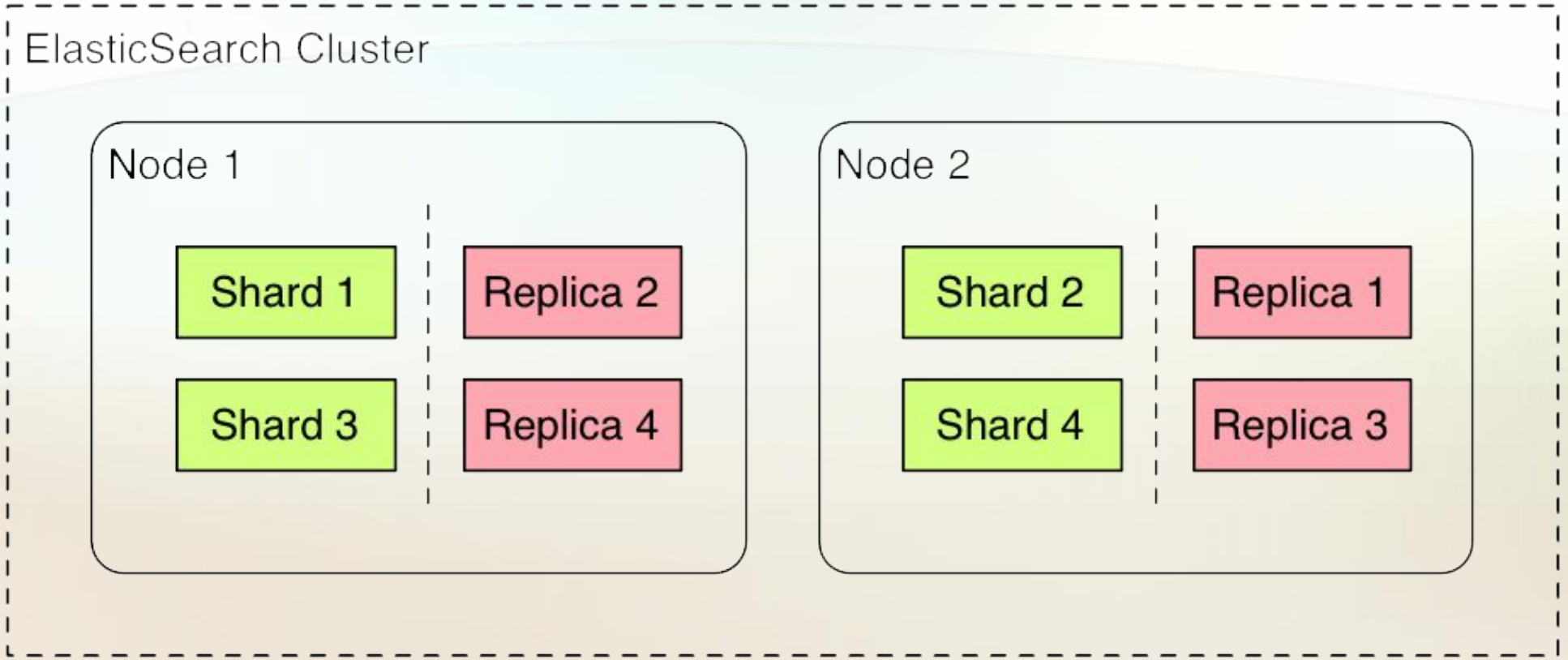
"log-level" => "INFO",

"class" => "MySecretApp.com.Transaction.Manager“ }

2.16 Elasticsearch

- Được phát triển bởi Shay Banon và dựa trên Apache Lucene.
- Là phần mềm open-source, theo giấy phép Apache License
- Phát triển bằng ngôn ngữ Java
- Sử dụng JSON cho việc truy xuất dữ liệu.
- Tìm kiếm, đánh chỉ mục và lưu trữ dữ liệu mạnh mẽ.
- Hoạt động theo cơ chế RESTful API
- Lưu trữ thành cụm, phân tán – cluster, shard & replica

2.17 Elasticsearch

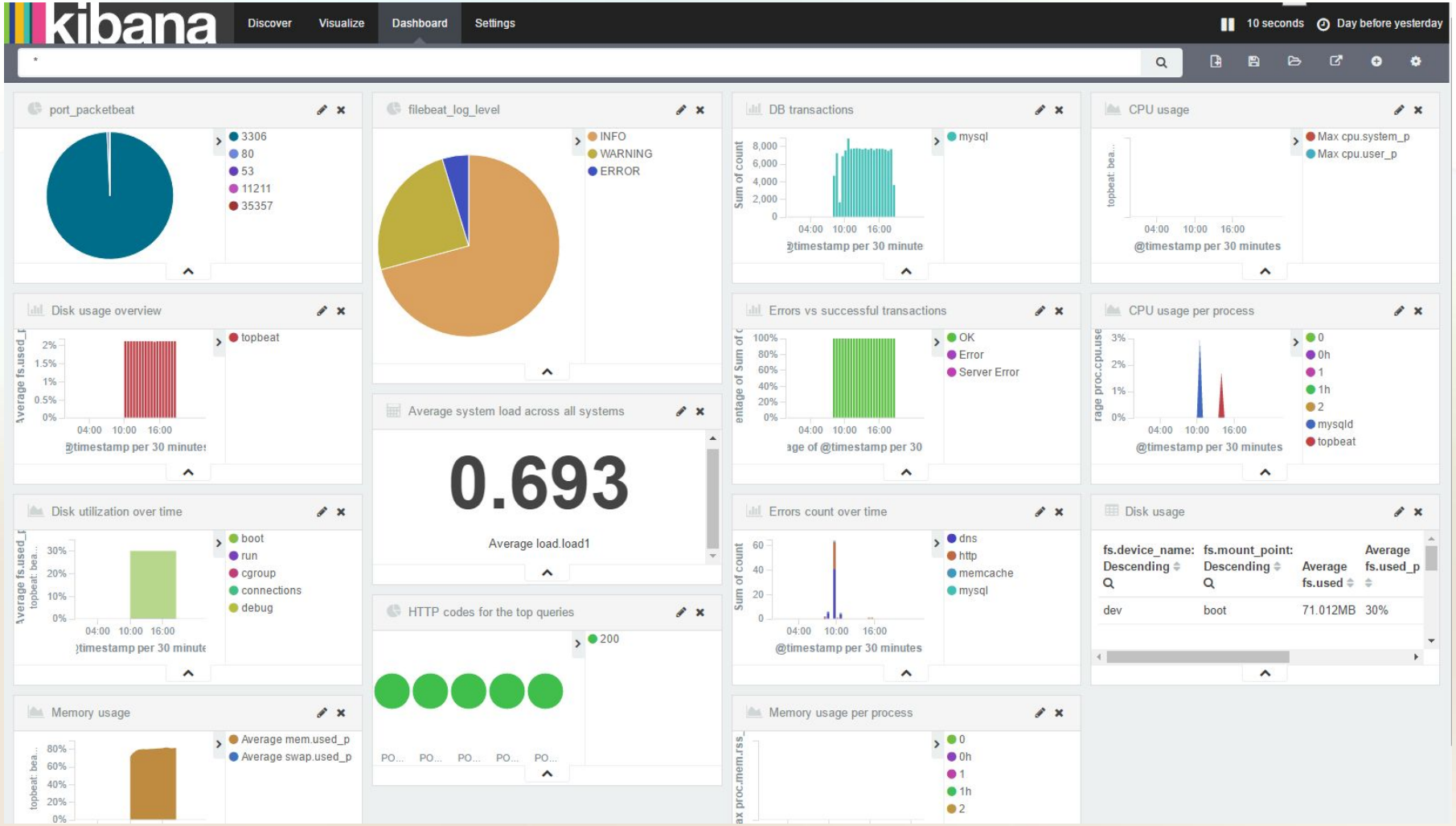


Source:https://goo.gl/1Ce99a

2.18 Kibana

- Cung cấp giao diện người dùng
- Phân tích cú pháp tìm kiếm
- Cung cấp nhiều dạng biểu đồ hiển thị dữ liệu

2.19 Kibana


kibana
Discover
Visualize
Dashboard
Settings
10 seconds
Day before yesterday
port_packetbeat
3306
80
53
11211
35357
filebeat_log_level
INFO
WARNING
ERROR
DB transactions
mysql
CPU usage
Max cpu.system_p
Max cpu.user_p
Disk usage overview
topbeat
Average system load across all systems
0.693
Average load.load1
Errors vs successful transactions
OK
Error
Server Error
CPU usage per process
0
0h
1
1h
2
mysqld
topbeat
Disk utilization over time
boot
run
cgroup
connections
debug
HTTP codes for the top queries
200
Errors count over time
dns
http
memcache
mysql
Disk usage
fs.device_name: Descending
fs.mount_point: Descending
Average fs.used
Average fs.used_p
dev
boot
71.012MB
30%
Memory usage
Average mem.used_p
Average swap.used_p
Memory usage per process

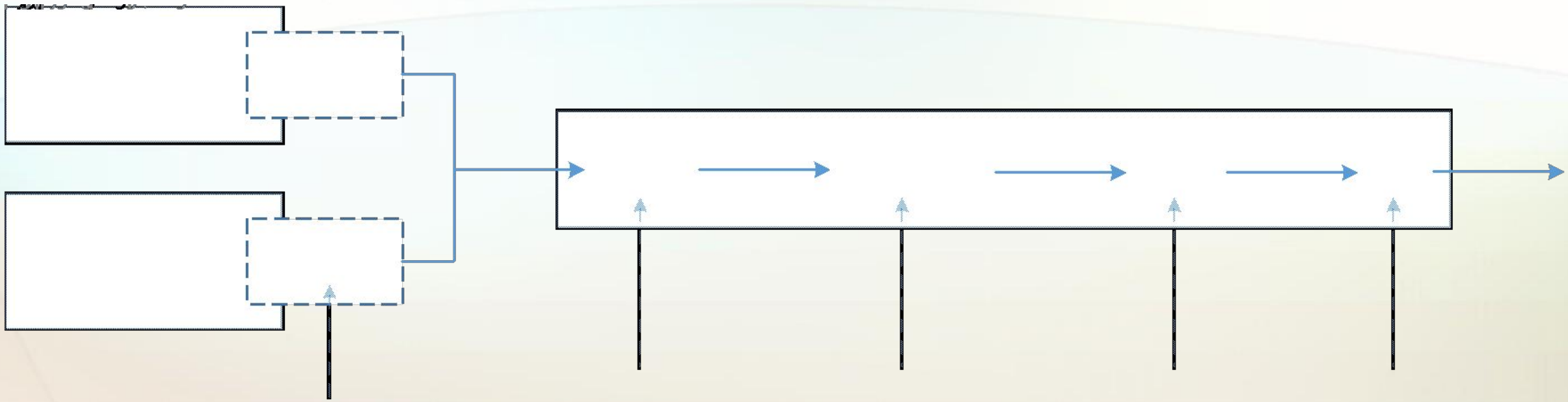
NỘI DUNG

1. Tìm hiểu về LOG
2. Tìm hiểu về ELK stack
3. Triển khai thử nghiệm
4. Thảo luận

3.1 Chuẩn bị

- 01 ELK server + 01 client server
- Hệ điều hành Ubuntu 14.04 64bit
- Elasticsearch version 2.4.1
- Logstash version 2.3.4
- Kibana version 4.5.4
- Filebeat version 1.3.1

3.2 Mô hình thử nghiệm



3.3 Kịch bản thử nghiệm

- Cài đặt ELK server theo script (tự viết).
- Tính năng thu thập dữ liệu
- Tính năng phân tích dữ liệu
- Tính năng tìm kiếm dữ liệu
- Tính năng hiển thị dữ liệu dạng biểu đồ

NỘI DUNG

1. Tìm hiểu về LOG
2. Tìm hiểu về ELK stack
3. Triển khai thử nghiệm
4. Thảo luận

Thảo luận

Source:https://goo.gl/Ek2TCj