How would you ship logs from an application running on an EC2 instance to ELK?

🔷 **Problem**
You have an **application running on an EC2 instance** (say, a Node.js app). It produces logs in /var/log/app/app.log.
👉 You want those logs to appear in **Kibana (via ELK)**.

🔷 **Solution: Use Filebeat → Logstash → Elasticsearch → Kibana**
Here's the **flow**:
[App on EC2] --> [Filebeat] --> [Logstash] --> [Elasticsearch] --> [Kibana]

🔷 **Step-by-Step**

1️⃣ Install Filebeat on the EC2 instance

- Filebeat is the **agent** that will read your log files.
- Install it using package manager (e.g., yum or apt).

Example (Amazon Linux):

```
sudo rpm --import https://artifacts.elastic.co/GPG-KEY-elasticsearch
sudo tee /etc/yum.repos.d/elastic.repo <<EOF
[elastic-7.x]
name=Elastic repository for 7.x packages
baseurl=https://artifacts.elastic.co/packages/7.x/yum
gpgcheck=1
enabled=1
autorefresh=1
type=rpm-md
EOF

sudo yum install filebeat -y
```

2️⃣ Configure Filebeat to watch the application log
Edit /etc/filebeat/filebeat.yml:

```
filebeat.inputs:
```

```
  - type: log
    enabled: true
    paths:
      - /var/log/app/app.log   # Your app log file

output.logstash:
  hosts: ["logstash-server:5044"]   # Replace with your Logstash server
```

👉 If you don’t need Logstash, you can directly output to Elasticsearch:

```
output.elasticsearch:
  hosts: ["http://elasticsearch-server:9200"]
```

3️⃣ Start Filebeat

```
sudo systemctl enable filebeat
sudo systemctl start filebeat
```

4️⃣ Logstash Configuration (Optional, for parsing/filters)

On your Logstash server, configure a pipeline: /etc/logstash/conf.d/app.conf

```
input {
  beats {
    port => 5044
  }
}

filter {
  grok {
    match => { "message" => "%{COMMONAPACHELOG}" }
  }
}

output {
  elasticsearch {
    hosts => ["http://elasticsearch-server:9200"]
    index => "app-logs-%{+YYYY.MM.dd}"
  }
}
```

5️⃣ Check in Kibana

- Go to **Kibana → Discover**.
- Choose the app-logs-* index pattern.
- You should now see your app logs live in Kibana. 🎉

🔷 **Analogy**

Think of this like:

- **Filebeat** = courier picking up letters (logs) from your EC2 house.
- **Logstash** = post office that sorts/parses letters.
- **Elasticsearch** = central storage warehouse.
- **Kibana** = shop window where you can browse all letters neatly.

🔷 **Real DevOps Use Case**

- **E-commerce app logs on EC2** → Filebeat ships logs → Elasticsearch → Kibana dashboards show real-time orders & errors.
- **Microservices on multiple EC2s** → all logs shipped to a **central ELK** → DevOps/SRE team troubleshoots faster.

What is the primary role of **Filebeat** in the ELK stack?
A) Store logs in Elasticsearch
B) Collect and ship logs from EC2 to Logstash/Elasticsearch
C) Parse and transform logs
D) Visualize logs in Kibana
Answer: B) Collect and ship logs from EC2 to Logstash/Elasticsearch

---

If you don't want to use Logstash, Filebeat can directly send logs to:
A) Kibana
B) Elasticsearch
C) Prometheus
D) Grafana
Answer: B) Elasticsearch

Which command ensures Filebeat starts automatically after a system reboot?
A) sudo yum install filebeat -y
B) sudo systemctl enable filebeat
C) sudo systemctl restart filebeat
D) sudo systemctl status filebeat
Answer: B) sudo systemctl enable filebeat

In the Logstash configuration, which plugin is commonly used to parse log formats like Apache logs?
A) mutate
B) grok
C) beats
D) date
Answer: B) grok

Which ELK component acts as the **central storage warehouse** for logs?
A) Kibana
B) Elasticsearch
C) Filebeat
D) Logstash
Answer: B) Elasticsearch

After setting up the pipeline, where would you verify if the logs are visible in Kibana?
A) Kibana → Dashboard
B) Kibana → Dev Tools
C) Kibana → Discover
D) Kibana → Alerts
Answer: C) Kibana → Discover