

CSE-4E3 Designing Data Intensive Applications

Project: Weather Stations Monitoring

Team Members

AbdElaziz Mohamed AbdElaziz	19015941
Veronica Romany Hanna	19016156
Mark Ehab Latif	19016213

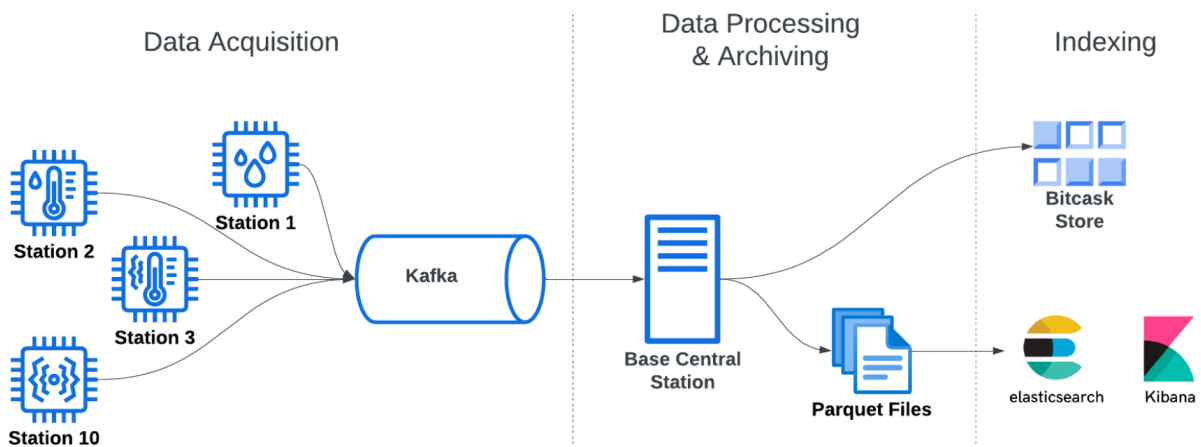
Problem Statement

The Internet of Things (IoT) is an important source of data streams in the modern digital world. The “Things” are huge in count and emit messages in very high frequency which flood the global internet. Hence, efficient stream processing is inevitable.

One use case is the distributed weather stations use case. Each “weather station” emits readings for the current weather status to the “central base station” for persistence and analysis. In this project, you will be required to implement the architecture of a weather monitoring system.

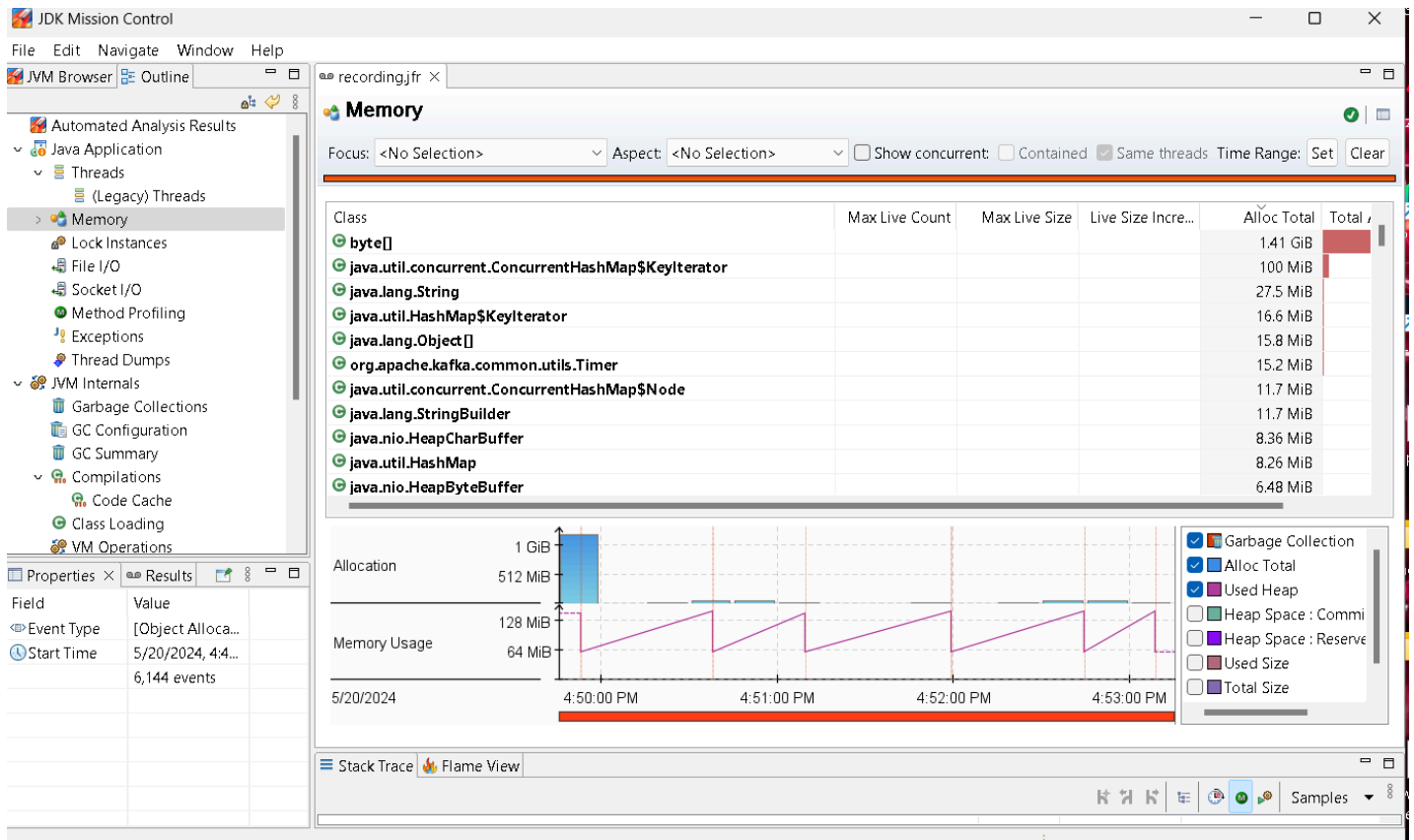
The system is composed of three stages:

- **Data Acquisition:** multiple weather stations that feed a queueing service (Kafka) with their readings.
- **Data Processing & Archiving:** The base central station is consuming the streamed data and archiving all data in the form of Parquet files.
- **Indexing:** two variants of index are maintained
 - Key-value store (Bitcask) for the latest reading from each individual station.
 - ElasticSearch / Kibana that are running over the Parquet files.

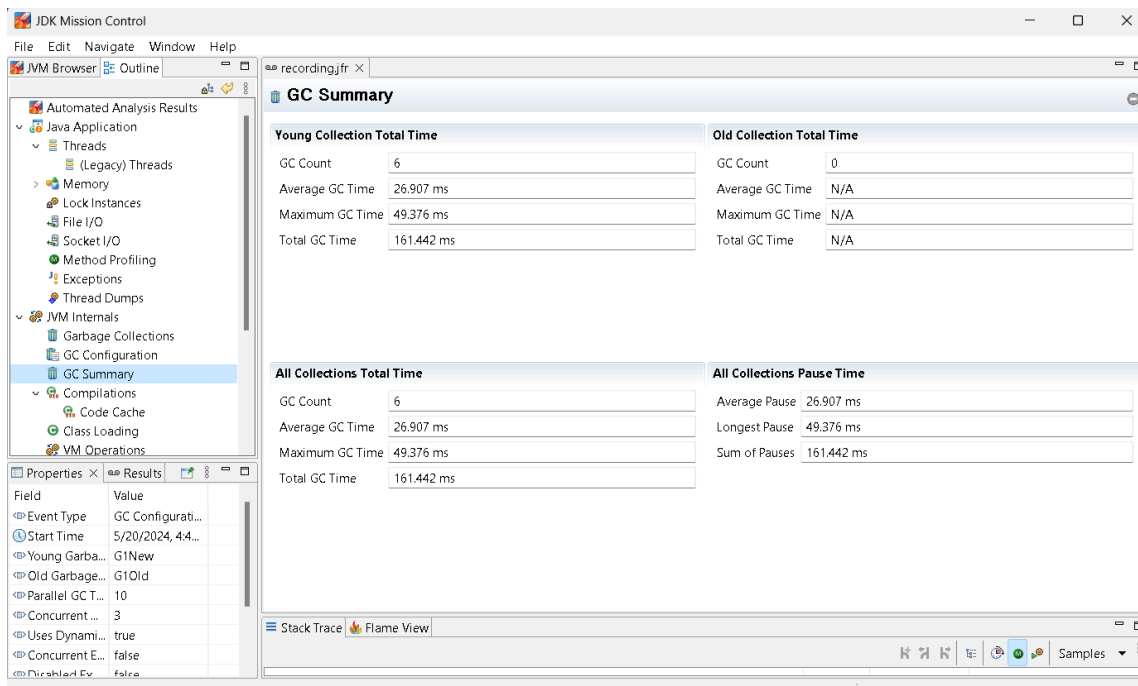
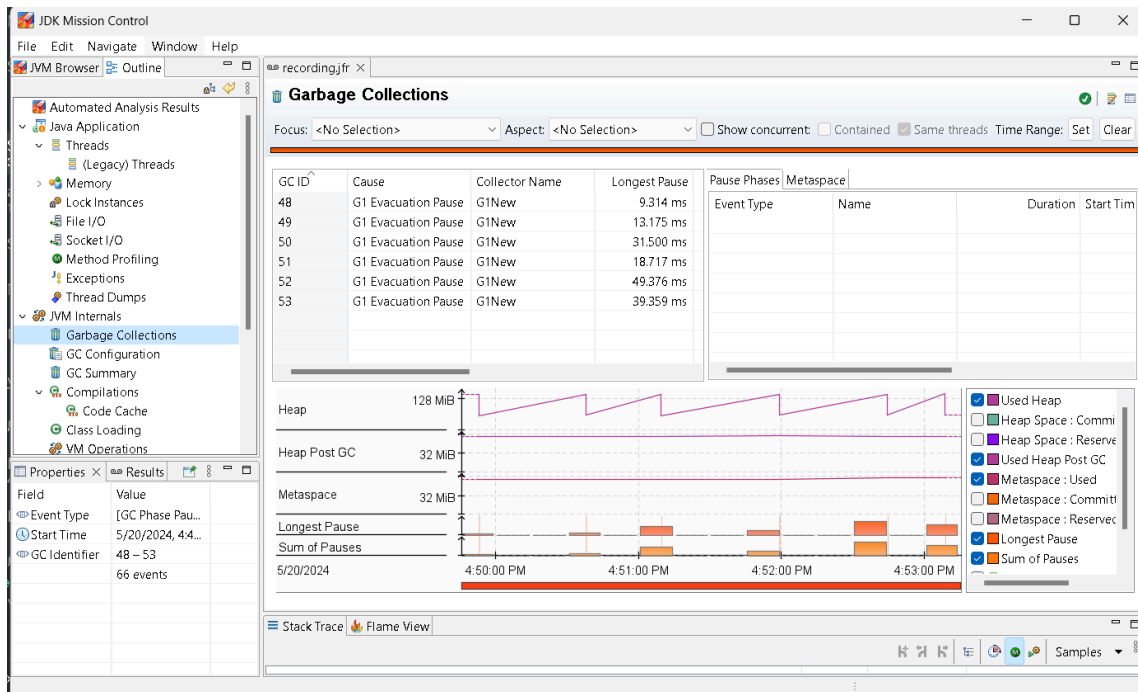


Profile Central Station using JFR

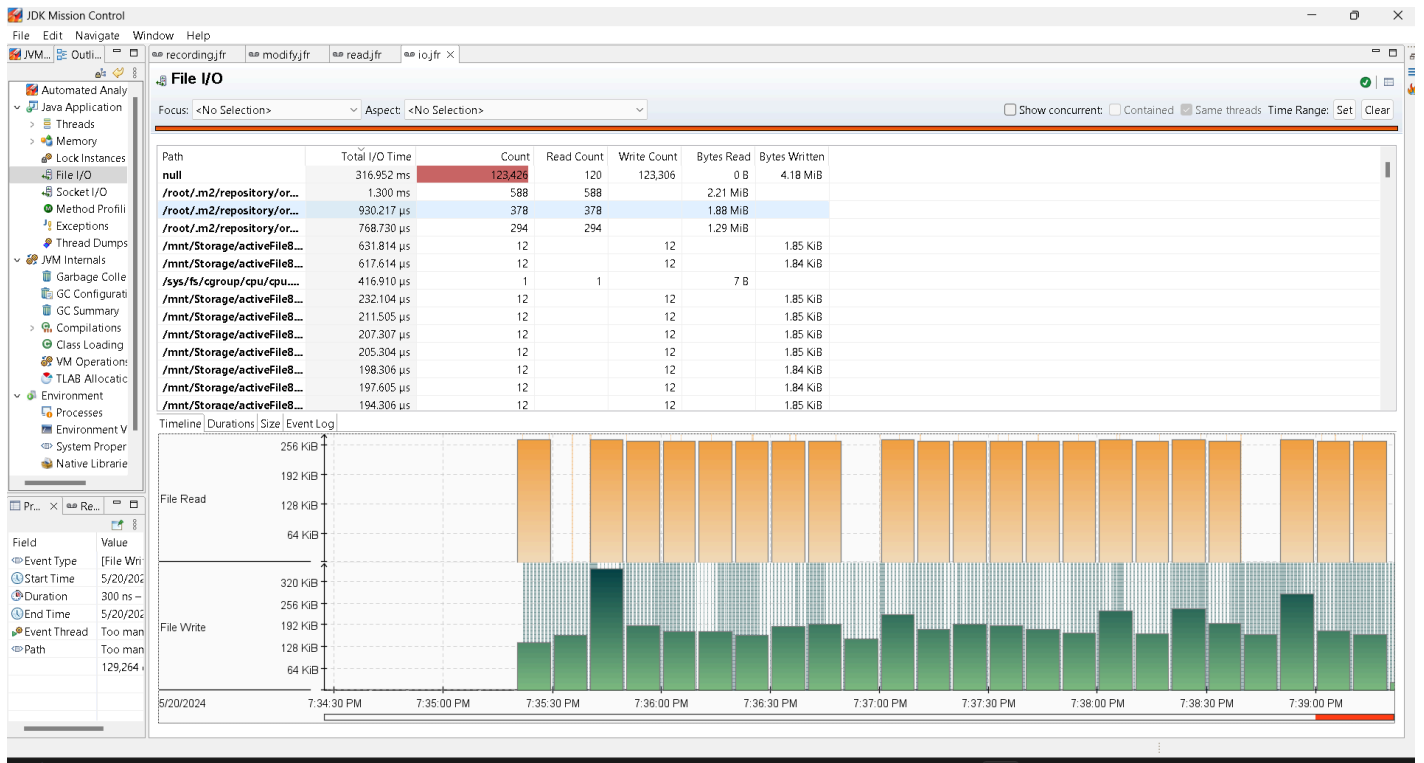
- Top 10 Classes with highest total memory



- GC pauses count
- GC maximum pause duration



- List of I/O operations



Docker & Kubernetes files

- K8s yaml file
 - Station 1

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: station1
  labels:
    app: station1
spec:
  replicas: 1
  selector:
    matchLabels:
      app: station1
  template:
    metadata:
      labels:
        app: station1
    spec:
      containers:
        - name: station1
          image: abdelaziz89/weather-station:v2.0
          ports:
            - containerPort: 8080
          env:
            - name: StationId
              value: "1"
            - name: latitude
              value: "31.2018"
            - name: longitude
              value: "29.9158"
```

- Kafka Processor

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: kafkaprocessor
  labels:
    app: kafkaprocessor
spec:
  replicas: 1
  selector:
    matchLabels:
      app: kafkaprocessor
  template:
    metadata:
      labels:
        app: kafkaprocessor
    spec:
      containers:
        - name: kafkaprocessor
          image: abdelaziz89/kafkaprocessor:v3.0
          ports:
            - containerPort: 8080
```

- Central Server

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: centralserver
  labels:
    app: centralserver
spec:
  replicas: 1
  selector:
    matchLabels:
      app: centralserver
  template:
    metadata:
      labels:
        app: centralserver
    spec:
      containers:
        - name: newserver
          image: abdelaziz89/central-server:v21.0
          ports:
            - containerPort: 8080
          volumeMounts:
            - mountPath: "/mnt/Volume"
              name: parquet
            - mountPath: "/mnt/Storage"
              name: parquet
      volumes:
        - name: parquet
          persistentVolumeClaim:
            claimName: serverstorage
```


- Persistent Volume

```
apiVersion: v1
kind: PersistentVolume
metadata:
  name: storage
spec:
  storageClassName: manual
  capacity:
    storage: 2Gi
  accessModes:
    - ReadWriteOnce
  hostPath:
    path: "/data"
```

- Persistent Volume Claim

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: serverstorage
spec:
  accessModes:
    - ReadWriteOnce
  storageClassName: manual
  resources:
    requests:
      storage: 2Gi
  volumeName: storage
```

- kibana

```

  apiVersion: apps/v1
  kind: Deployment
  metadata:
    name: kibana-deployment
    labels:
      app: kibana
  spec:
    replicas: 1
    selector:
      matchLabels:
        app: kibana
    template:
      metadata:
        labels:
          app: kibana
      spec:
        containers:
          - name: kibana
            image: kibana:8.4.0
            ports:
              - containerPort: 5601
            env:
              - name: ELASTICSEARCH_HOSTS
                value: http://elasticsearch-service:9200
---
  apiVersion: v1
  kind: Service
  metadata:
    name: kibana-service
    labels:
      app: kibana
  spec:
    ports:
      - port: 5601
        targetPort: 5601
    selector:
      app: kibana
```

- Elastic Search

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: elasticsearch-deployment
  labels:
    app: elasticsearch
spec:
  replicas: 1
  selector:
    matchLabels:
      app: elasticsearch
  template:
    metadata:
      labels:
        app: elasticsearch
    spec:
      containers:
        - name: elasticsearch
          image: elasticsearch:8.4.0
          ports:
            - containerPort: 9200
          env:
            - name: xpack.security.enabled
              value: "false"
            - name: discovery.type
              value: single-node
            - name: ES_JAVA_OPTS
              value: "-Xms512m -Xmx512m"
---
apiVersion: v1
kind: Service
metadata:
  name: elasticsearch-service
  labels:
    app: elasticsearch
spec:
  ports:
    - port: 9200
      targetPort: 9200
  selector:
    app: elasticsearch
```

- Weather Station Dockerfile
- Central Server Dockerfile

```
FROM maven:3.9.2-eclipse-temurin-17-alpine

WORKDIR /CentralServer
EXPOSE 8080
# Copy the Maven project files into the Docker container
COPY src ./src
COPY pom.xml .

# Copy the Avro schema file into the Docker container
COPY src/main/java/BaseCentralStation/weather_record.avsc
/CentralServer/src/main/java/BaseCentralStation/weather_record.avsc

# Build the Maven project inside the Docker container
RUN mvn clean install

# Specify the command to run your Java application
CMD ["mvn", "exec:java"]
```

Kubernetes GUI

The screenshot shows the Kubernetes GUI with the 'Deployments' tab selected. The left sidebar contains a navigation menu with options like Cluster, Applications, Nodes, Workloads, Overview, Pods, Daemon Sets, Stateful Sets, Replica Sets, Replication Controllers, Jobs, Cron Jobs, Config, Network, Services, Endpoints, Ingresses, Ingress Classes, Network Policies, Port Forwarding, and Storage. The main panel displays a table of 15 Deployments across various namespaces.

Name	Namespace	Pods	Replicas	Age	Conditions
centralserver	default	1/1	1	144m	Available Progressing
elasticsearch-deployment	default	1/1	1	11h	Available Progressing
kafkaprocessor	default	1/1	1	10h	Available Progressing
kibana-deployment	default	1/1	1	10h	Available Progressing
station1	default	1/1	1	147m	Available Progressing
station10	default	1/1	1	10m	Available Progressing
station2	default	1/1	1	99m	Available Progressing
station3	default	1/1	1	74m	Available Progressing
station4	default	1/1	1	64m	Available Progressing
station5	default	1/1	1	30m	Available Progressing
station6	default	1/1	1	25m	Available Progressing
station7	default	1/1	1	21m	Available Progressing
station8	default	1/1	1	17m	Available Progressing
station9	default	1/1	1	14m	Available Progressing
coredns	kube-system	1/1	1	11h	Available Progressing

The screenshot shows the Kubernetes GUI with the 'Services' tab selected. The left sidebar is the same as the previous screenshot. The main panel displays a table of 8 Services across various namespaces.

Name	Namespace	Type	Cluster IP	Ports	External IP	Selector	Age	Status
elasticsearch-service	default	ClusterIP	10.110.243.2	9200/TCP	-	app=elastic	11h	Active
kafka-release	kafka	ClusterIP	10.110.211.2	9092:kafka-dien	-	app.kubern	11h	Active
kafka-release-headless	kafka	ClusterIP	None	9092:kafka-dien	-	app.kubern	11h	Active
kafka-release-zookeepe	kafka	ClusterIP	10.100.150.8	2181:dien/TCP	-	app.kubern	11h	Active
kafka-release-zookeepe	kafka	ClusterIP	None	2181:dien/TCP	-	app.kubern	11h	Active
kibana-service	default	ClusterIP	10.105.244.2	5601/TCP	-	app=kibana	10h	Active
kube-dns	kube-system	ClusterIP	10.96.0.10	53/UDP, 53/TCP	-	k8s-app=ku	11h	Active
kubernetes	default	ClusterIP	10.96.0.1	443:8443/TCP	-		11h	Active

The screenshot shows the Kubernetes dashboard interface. On the left, a sidebar menu lists various resources, with 'Storage' expanded and 'Persistent Volume Claims' selected. The main panel displays the 'Persistent Volume Claims' tab, showing a table of 10 items. The table columns are Name, Namespace, Storage class, Size, Pods, Age, and Status. All claims are in a 'Bound' status.

Name	Namespace	Storage class	Size	Pods	Age	Status
data-kafka-release-0	kafka	standard	8Gi	kafka-rele:	11h	Bound
data-kafka-release-1	kafka	standard	8Gi	kafka-rele:	11h	Bound
data-kafka-release-2	kafka	standard	8Gi	kafka-rele:	11h	Bound
data-kafka-release-zookeeper-0	kafka	standard	2Gi	kafka-rele:	11h	Bound
data-kafka-release-zookeeper-1	kafka	standard	2Gi	kafka-rele:	11h	Bound
data-kafka-release-zookeeper-2	kafka	standard	2Gi	kafka-rele:	11h	Bound
logs-kafka-release-0	kafka	standard	8Gi	kafka-rele:	11h	Bound
logs-kafka-release-1	kafka	standard	8Gi	kafka-rele:	11h	Bound
logs-kafka-release-2	kafka	standard	8Gi	kafka-rele:	11h	Bound
serverstorage	default	manual	2Gi	centralser	10h	Bound

The screenshot shows the 'Persistent Volumes' tab in the Kubernetes dashboard. It displays a table of 10 items. The table columns are Name, Storage Class, Capacity, Claim, Age, and Status. Most volumes are of 'standard' storage class and are 'Bound' to a claim.

Name	Storage Class	Capacity	Claim	Age	Status
pvc-06b8150f-f77c-409a-a00c-c320139a93d5	standard	8Gi	logs-kafka-	11h	Bound
pvc-2ae0445d-4eaa-4f18-863d-635b074c4df1	standard	8Gi	logs-kafka-	11h	Bound
pvc-2b90e63b-18d4-4d9e-83a8-fff178527058	standard	8Gi	data-kafka	11h	Bound
pvc-746af3ec-b89b-45b4-8fb1-a293a0add810	standard	2Gi	data-kafka	11h	Bound
pvc-7487fbc9-9f0a-4528-9a32-4eb5c2a199da	standard	8Gi	data-kafka	11h	Bound
pvc-98092cfd-cc62-4e2a-8907-05909300d941	standard	2Gi	data-kafka	11h	Bound
pvc-c951b2e9-1617-4183-aaa8-3349f171efdb	standard	8Gi	data-kafka	11h	Bound
pvc-ca291566-0b4f-410d-b16a-c61e80f3743b	standard	8Gi	logs-kafka-	11h	Bound
pvc-f0463558-ce88-4f11-a438-4cc47b4e1d1a	standard	2Gi	data-kafka	11h	Bound
storage	manual	2Gi	serverstor	10h	Bound

Name	Namespace	Contain...	CPU	Memory	Restarts	Controlled By	Node	QoS	Age	Status
kube-apiserver-minikube	kube-system		N/A	N/A	3	Node	minikube	Burstable	11h	Running
kube-controller-manager-minikube	kube-system		N/A	N/A	2	Node	minikube	Burstable	11h	Running
kube-proxy-4c2vb	kube-system		N/A	N/A	2	DaemonSet	minikube	BestEffort	11h	Running
kube-scheduler-minikube	kube-system		N/A	N/A	2	Node	minikube	Burstable	11h	Running
station1-84888fd4f6-c25nm	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	144m	Running
station10-7fb5c48d69-jxpkb	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	6m29s	Running
station2-7f54598df5-hv2dv	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	95m	Running
station3-685bfb88f-mvbc6	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	71m	Running
station4-5dd45645b-9hpx2	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	60m	Running
station5-75775f4dfc-sl8ds	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	26m	Running
station6-56f768b8c-9hd18	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	21m	Running
station7-586f9f64-sfphh	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	17m	Running
station8-7d76959b9d-xq59h	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	13m	Running
station9-5f85d76cd5-zpzs1	default		N/A	N/A	0	ReplicaSet	minikube	BestEffort	11m	Running
storage-provisioner	kube-system		N/A	N/A	10		minikube	BestEffort	11h	Running

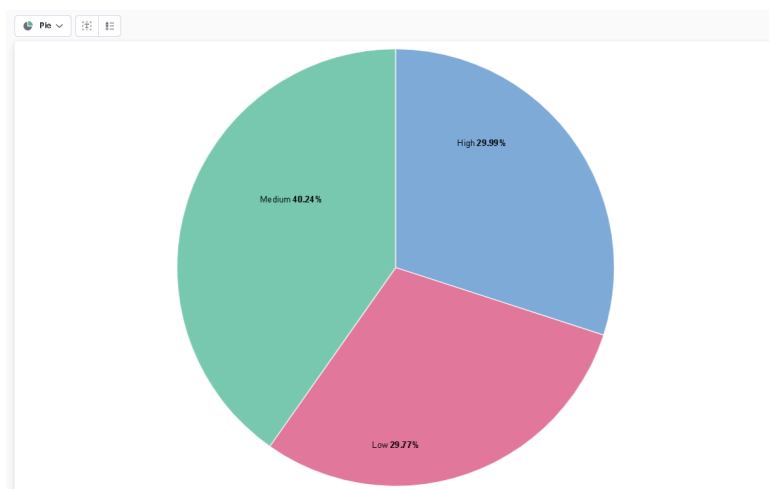
```

Downloaded from central: https://repo.maven.apache.org/maven2/org/ow2/asm/asm-tree/9.6/asm-tree-9.6.jar (52 kB at 22 kB/s)
Downloaded from central: https://repo.maven.apache.org/maven2/org/codehaus/plexus/plexus-utils/4.0.0/plexus-utils-4.0.0.jar (192 kB at 66 kB/s)
Downloaded from central: https://repo.maven.apache.org/maven2/org/ow2/asm/asm/9.6/asm-9.6.jar (124 kB at 44 kB/s)
log4j:WARN No appenders could be found for logger (org.apache.kafka.clients.producer.ProducerConfig).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
Raining detected: {"station_id":1,"s_no":21,"weather":{"temperature":74,"humidity":71,"wind_speed":16},"status_timestamp":1716224400,"battery_status":"Medium"}
Raining detected: {"station_id":1,"s_no":22,"weather":{"temperature":73,"humidity":75,"wind_speed":15},"status_timestamp":1716228000,"battery_status":"Medium"}
Raining detected: {"station_id":1,"s_no":23,"weather":{"temperature":73,"humidity":78,"wind_speed":16},"status_timestamp":1716231600,"battery_status":"Medium"}
Raining detected: {"station_id":1,"s_no":24,"weather":{"temperature":72,"humidity":79,"wind_speed":16},"status_timestamp":1716235200,"battery_status":"Low"}
Raining detected: {"station_id":10,"s_no":24,"weather":{"temperature":74,"humidity":76,"wind_speed":8},"status_timestamp":1716235200,"battery_status":"Low"}
Raining detected: {"station_id":1,"s_no":45,"weather":{"temperature":74,"humidity":71,"wind_speed":16},"status_timestamp":1716224400,"battery_status":"High"}
Raining detected: {"station_id":1,"s_no":46,"weather":{"temperature":73,"humidity":75,"wind_speed":15},"status_timestamp":1716228000,"battery_status":"Low"}
Raining detected: {"station_id":1,"s_no":47,"weather":{"temperature":73,"humidity":78,"wind_speed":16},"status_timestamp":1716231600,"battery_status":"Medium"}
Raining detected: {"station_id":1,"s_no":48,"weather":{"temperature":72,"humidity":79,"wind_speed":16},"status_timestamp":1716235200,"battery_status":"High"}
Raining detected: {"station_id":10,"s_no":48,"weather":{"temperature":74,"humidity":76,"wind_speed":8},"status_timestamp":1716235200,"battery_status":"Low"}
Raining detected: {"station_id":1,"s_no":69,"weather":{"temperature":74,"humidity":71,"wind_speed":16},"status_timestamp":1716224400,"battery_status":"Medium"}
Raining detected: {"station_id":1,"s_no":70,"weather":{"temperature":73,"humidity":75,"wind_speed":15},"status_timestamp":1716228000,"battery_status":"High"}
Raining detected: {"station_id":1,"s_no":71,"weather":{"temperature":73,"humidity":78,"wind_speed":16},"status_timestamp":1716231600,"battery_status":"High"}
Raining detected: {"station_id":1,"s_no":72,"weather":{"temperature":72,"humidity":79,"wind_speed":16},"status_timestamp":1716235200,"battery_status":"Medium"}
Raining detected: {"station_id":10,"s_no":72,"weather":{"temperature":74,"humidity":76,"wind_speed":8},"status_timestamp":1716235200,"battery_status":"High"}
Raining detected: {"station_id":1,"s_no":93,"weather":{"temperature":74,"humidity":71,"wind_speed":16},"status_timestamp":1716224400,"battery_status":"Low"}
Raining detected: {"station_id":1,"s_no":94,"weather":{"temperature":73,"humidity":75,"wind_speed":15},"status_timestamp":1716228000,"battery_status":"Medium"}
Raining detected: {"station_id":1,"s_no":95,"weather":{"temperature":73,"humidity":78,"wind_speed":16},"status_timestamp":1716231600,"battery_status":"High"}
Raining detected: {"station_id":1,"s_no":96,"weather":{"temperature":72,"humidity":79,"wind_speed":16},"status_timestamp":1716235200,"battery_status":"High"}
Raining detected: {"station_id":10,"s_no":96,"weather":{"temperature":74,"humidity":76,"wind_speed":8},"status_timestamp":1716235200,"battery_status":"Low"}
Raining detected: {"station_id":1,"s_no":117,"weather":{"temperature":74,"humidity":71,"wind_speed":16},"status_timestamp":1716224400,"battery_status":"High"}

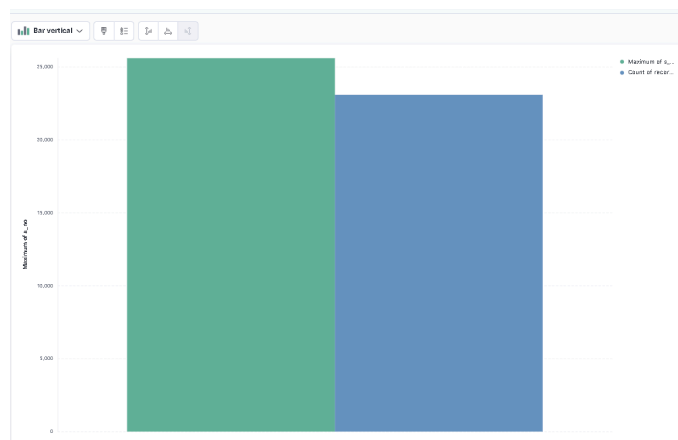
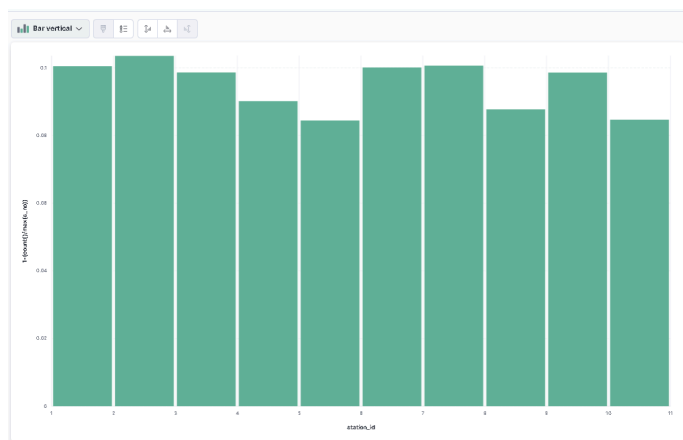
```

Kibana visualization confirming

- Battery status distribution (30% low - 40% medium - 30% high)



- 10% dropped messages



Sample Parquet File

- Parquet File Directory

station5			33 minutes ago	drwxr-xr-x
station6			28 minutes ago	drwxr-xr-x
20240520			4 minutes ago	drwxr-xr-x
.Project_1716211048062.parquet.crc	32 Bytes		28 minutes ago	-rw-r--r--
.Project_1716211294088.parquet.crc	40 Bytes		23 minutes ago	-rw-r--r--
.Project_1716211892901.parquet.crc	52 Bytes		13 minutes ago	-rw-r--r--
.Project_1716212072933.parquet.crc	36 Bytes		11 minutes ago	-rw-r--r--
.Project_1716212262086.parquet.crc	36 Bytes		7 minutes ago	-rw-r--r--
.Project_1716212273098.parquet.crc	24 Bytes		7 minutes ago	-rw-r--r--
.Project_1716212418981.parquet.crc	32 Bytes		5 minutes ago	-rw-r--r--
.Project_1716212464483.parquet.crc	28 Bytes		4 minutes ago	-rw-r--r--
Project_1716211048062.parquet	2.6 kB		28 minutes ago	-rw-r--r--
Project_1716211294088.parquet	3.5 kB		23 minutes ago	-rw-r--r--
Project_1716211892901.parquet	5.4 kB		13 minutes ago	-rw-r--r--
Project_1716212072933.parquet	3.1 kB		11 minutes ago	-rw-r--r--
Project_1716212262086.parquet	3.2 kB		7 minutes ago	-rw-r--r--
Project_1716212273098.parquet	1.9 kB		7 minutes ago	-rw-r--r--

Containers	minikube				STATUS
Images	get.io/k8s-minikube/kicbasev0.0.43				Running (2 hours ago)
Volumes	359107c4b6f4				
Builds	Logs Inspect Bind mounts Exec Files Stats				Open file editor
Dev Environments					
Docker Scout					
Extensions					
Add Extensions					
	Name	Note	Size	Last modified	Mode
	station1			2 hours ago	drwxr-xr-x
	station10			11 minutes ago	drwxr-xr-x
	station2			2 hours ago	drwxr-xr-x
	station3			1 hour ago	drwxr-xr-x
	station4			1 hour ago	drwxr-xr-x
	station5			33 minutes ago	drwxr-xr-x
	station6			27 minutes ago	drwxr-xr-x
	station7			22 minutes ago	drwxr-xr-x
	station8			20 minutes ago	drwxr-xr-x
	station9			16 minutes ago	drwxr-xr-x
	dev			4 hours ago	drwxr-xr-x
	docker.key		3.7 kB	1 month ago	-rw-r--r--

- Parquet File Sample

[Project_1716203944681.parquet](#)

station_id - i32	s_no - i32	weather - struct[3]	status_timestamp - i64	battery_status - str
1	1	{82,44,25}	1716152400	Medium
1	2	{82,45,26}	1716156000	Low
1	3	{82,45,26}	1716159600	Medium
1	4	{82,44,24}	1716163200	Low
1	5	{82,48,22}	1716166800	Low
1	6	{82,49,20}	1716170400	Medium
1	7	{82,49,16}	1716174000	Medium
1	8	{82,53,18}	1716177600	High
1	9	{80,57,23}	1716181200	Low
1	10	{78,51,15}	1716184800	Medium
1	11	{79,49,13}	1716188400	Medium
1	12	{83,39,11}	1716192000	High
1	13	{89,23,16}	1716195600	Medium
1	14	{81,43,23}	1716199200	Medium
1	15	{78,52,23}	1716202800	Medium
1	16	{78,51,21}	1716206400	High
1	17	{78,52,20}	1716210000	High
1	18	{77,55,19}	1716213600	Medium

Sample BitCask Riak LSM directory

- Files in directory

Name	Note	Size	Last modified	Mode
.dockerenv		0 Bytes	11 hours ago	-rwxr-xr-x
bin -> usr/bin		7 Bytes	3 months ago	lrwxrwxrwx
boot			2 years ago	drwxr-xr-x
CHANGELOG		7.6 kB	1 month ago	-rw-r--r--
data	ADDED		1 hour ago	drwxr-xr-x
activeFile-144		1.5 kB	43 seconds ago	-rw-r--r--
activeFile-144.copy		1.5 kB	43 seconds ago	-rw-r--r--
activeFile-144.hint		251 Bytes	43 seconds ago	-rw-r--r--
activeFile-2294		1.9 kB	42 seconds ago	-rw-r--r--
activeFile-2294.copy		1.9 kB	42 seconds ago	-rw-r--r--
activeFile-2294.hint		301 Bytes	42 seconds ago	-rw-r--r--
activeFile-2295		1.9 kB	41 seconds ago	-rw-r--r--
activeFile-2295.copy		1.9 kB	41 seconds ago	-rw-r--r--
activeFile-2295.hint		301 Bytes	41 seconds ago	-rw-r--r--
activeFile-2296		1.9 kB	39 seconds ago	-rw-r--r--
activeFile-2296.copy		1.9 kB	39 seconds ago	-rw-r--r--
activeFile-2296.hint		300 Bytes	39 seconds ago	-rw-r--r--

- BitCask In Memory Key Value Directory Sample

```
{ "station_id": 4, "s_no": 4173, "weather": { "temperature": 88, "humidity": 27, "wind_speed": 25, "status_timestamp": 1716224400, "battery_status": "High" } }
Successfully Added this Entry To BitCask File >>
Successfully Added this Entry To In Memory directory >>
Successfully Added this Entry To Hint File >>
Key Directory Contents:
Key: [49]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=792, valueSize=140, timeStamp=1716212882286}
Key: [55]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=474, valueSize=140, timeStamp=1716212882174}
Key: [54]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2514', valuePosition=1592, valueSize=140, timeStamp=1716212881729}
Key: [57]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=315, valueSize=142, timeStamp=1716212882080}
Key: [56]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=0, valueSize=140, timeStamp=1716212881814}
Key: [49, 48]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2514', valuePosition=636, valueSize=143, timeStamp=1716212880942}
Key: [51]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=157, valueSize=141, timeStamp=1716212881929}
Key: [50]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=631, valueSize=144, timeStamp=1716212882249}
Key: [53]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2514', valuePosition=1749, valueSize=142, timeStamp=1716212881779}
Key: [52]
Value: KeyDirectoryEntry{fileId='/mnt/Storage/activeFile2515', valuePosition=949, valueSize=141, timeStamp=1716212882556}
End of Key Directory Contents
```

● BitCask File Entries

- Active File Entry

```
public class BitCaskEntry {  
    private long timeStamp;  
    private int keyLength;  
    private int valueLength ;  
    private byte [] key;  
    private byte [] value;  
}
```

- Hint File Entry

```
public class HintFileEntry {  
    private long tstamp;  
    private int keySize ;  
    private int valueSize ;  
    private long valuePosition;  
    private byte[] key;  
}
```

- Key Directory In Memory Entry

```
public class KeyDirectoryEntry {  
    private String fileId ;  
    private long valuePosition;  
    private int valueSize ;  
    private long timeStamp;  
}
```

Bonus

- Open-Meteo

```
public class OpenMeteoAPI {
    private static final String OPEN_METEO_API_URL = "https://api.open-meteo.com/weather";

    private final KafkaProducer<String, String> kafkaProducer;
    private final Random random = new Random();
    public OpenMeteoAPI(KafkaProducer<String, String> kafkaProducer) {
        this.kafkaProducer = kafkaProducer;
    }
    public int seq =1 ;

    public void fetchAndPublishWeatherData(int StationId,double latitude, double longitude) {
        try {
            while(true)
            {
                // Construct API URL with latitude, longitude, and other parameters
                String apiUrl = "https://api.open-meteo.com/v1/forecast?latitude=" +
                    latitude + "&longitude=" + longitude +
                    "&hourly=relativehumidity_2m,windspeed_80m,temperature_80m" +
                    "current_weather=true&temperature_unit=fahrenheit&timeformat=unixtime" +
                    "&forecast_days=1&timezone=Africa%2FCairo";

                // Create HTTP connection
                URL url = new URL(apiUrl);
                HttpURLConnection connection = (HttpURLConnection) url.openConnection();
                connection.setRequestMethod("GET");
                // Check response code
                int responseCode = connection.getResponseCode();
            }
        }
    }
}
```

• Enterprise Integration Patterns

1. Adapter

```
// Parse JSON response
JSONObject jsonObject = new JSONObject(response.toString());
System.out.println(jsonObject);
// Extract hourly data object
JSONObject hourlyData = jsonObject.getJSONObject("hourly");

// Extract arrays for each parameter
JSONArray relativeHumidityArray = hourlyData.getJSONArray("relativehumidity_2m");
JSONArray windSpeedArray = hourlyData.getJSONArray("windspeed_80m");
JSONArray temperatureArray = hourlyData.getJSONArray("temperature_80m");
JSONArray timeArray = hourlyData.getJSONArray("time");
```

2. Aggregator

```
if (recordCount >= BATCH_SIZE) {
    // Write the batch to a Parquet file
    try {
        writeRecordsToParquet(stationId, stationRecords, avroSchema);
    } catch (IOException e) {
        e.printStackTrace();
        // Optionally handle the exception, e.g., skip the batch
    } finally {
        stationRecords.clear();
        recordCount = 0;
    }
}
```

3. Pipe and Filter

```
private static boolean isRaining(String value) {
    try {
        // Parse the JSON message
        JSONObject json = new JSONObject(value);

        // Extract humidity from the weather data
        JSONObject weatherData = json.getJSONObject("weather");
        int humidity = weatherData.getInt("humidity");

        // Check if humidity is higher than 70%
        if (humidity > 70) {
            // Print value if it's raining
            System.out.println("Raining detected: " + value);
            String message = "Raining detected: " + value;
            producer.send(new ProducerRecord<>("processor", message));
            return true;
        }
    } catch (Exception e) {
        // Handle parsing or any other exceptions
        e.printStackTrace();
    }
    return false;
}
```

4. Envelope Wrapper

```
kafkaProducer.send(new ProducerRecord<>("Project", Integer.toString(StationId), kafkaMessage.toString()));
kafkaProducer.flush();
```

```
for (ConsumerRecord<String, String> record : records) {
    GenericRecord avroRecord = createAvroRecord(record.value(), avroSchema);
    recordCount++;
    System.out.println(record.value());
    int stationId = (int) avroRecord.get("station_id");
    List<GenericRecord> stationRecords = stationRecordsMap.get(stationId);
    stationRecords.add(avroRecord);
}
```

5. Polling Consumer

```
ConsumerRecords<String, String> records = consumer.poll(Duration.ofMillis(100));
```

6. Enricher Content

```
// Method to build weather message with specified attributes
public static JSONObject buildWeatherMessage(int stationId, int sNo, int relativeHumidity, int windSpeed, int
temperature, long time) {

    JSONObject kafkaMessage = new JSONObject();
    kafkaMessage.put("station_id", stationId);
    kafkaMessage.put("s_no", sNo);
    kafkaMessage.put("battery_status", getRandomBatteryStatus());
    kafkaMessage.put("status_timestamp", time);
    JSONObject weatherData = new JSONObject();
    weatherData.put("humidity", relativeHumidity);
    weatherData.put("temperature", temperature);
    weatherData.put("wind_speed", windSpeed);
    kafkaMessage.put("weather", weatherData);
    return kafkaMessage;
}
```

Github Repo

<https://github.com/Abdelaziz25/Weather-Stations-Monitoring>