

ASSIGNMENT 2 : Real-time Server Monitoring Distributed Pipeline with Apache Kafka and Spark

Team No : 126

Team Name : A2_Team_19_RRC

Date : 22/10/2025

Team member details :

NAME: DISHAN D (Team leader) (CONSUMER 1)

SRN: PES1UG23CS196

NAME: HARSHAA VARDHANA KV (CONSUMER 2)

SRN: PES1UG23CS236

NAME: NAGARJUN N H (PRODUCER)

SRN: PES1UG23CS375

NAME: GAGAN BHARADWAJ K (BROKER)

SRN: PES1UG23CS212

SCREENSHOT OF OUTPUTS :

1. Zero tier network screenshot

The screenshot shows the ZeroTier Central interface for a network named "team_126 network". The network ID is 68bea79acf7249fb. It displays 4 included devices. The table lists the following members:

| Address | Name/Desc | Managed IPs | Last Seen | Version | Physical IP | OS | Architecture |
|---------------------------------|-----------|---------------|-----------|---------|-----------------|----|--------------|
| 106FB9F0C6 fa:59:1d:76:6a:61 | | 172.24.63.248 | 2 minutes | 1.16.0 | 223.231.177.197 | | |
| 381DA4EB58 fa:71:6f:6b:71:ff | | 172.24.68.72 | 1 minute | 1.16.0 | 106.51.200.69 | | |
| 3EFFF365B fa:77:8d:90:acfc | | 172.24.52.130 | 1 minute | 1.16.0 | 106.51.202.15 | | |
| 42F24E843D fa:db:80:81:1e:9a | | 172.24.91.49 | 1 minute | 1.16.0 | 49.205.150.56 | | |

At the bottom, there are fields for "E-Mail Join instructions" (alice@example.com) and "Manually Add Member" (#####), along with "Add New Member" and "Show Applications" buttons.

Shows that all the 4 devices are connected to the zero tier network named – “team 126 network”

PORTAL OUTPUT SCREENSHOTS :

CONSUMER 1 :

Big Data Assignment 2 Evaluation

Team Number
126

Part
CPU_MEM

Result CSV
 team_126_CPU_MEM.csv

We got an accuracy of 100 % for the consumer 1's spark job which was on the topics – cpu and mem

CONSUMER 2:

Big Data Assignment 2 Evaluation

Team Number

Part

Result CSV

Choose File

Upload and Evaluate

Evaluation Results

Accuracy: 100% | Mismatches: 0 / 14400

| Submitted (Mismatched Rows) | Expected (Ground Truth) |
|--|--|
| Submitted (Mismatched Rows) | Expected (Ground Truth) |
| server_id window_start window_end max_net_in max_disk_io alert | server_id window_start window_end max_net_in max_disk_io alert |

We got an accuracy of 100 % for the consumer 2's spark job which was on the topics – net and disk

EXECUTION SCREENSHOTS :

1.PRODUCER

```
pes1ug23cs375@Ubuntu22:~/Assignment-2$ ping 172.24.63.248
PING 172.24.63.248 (172.24.63.248) 56(84) bytes of data.
64 bytes from 172.24.63.248: icmp_seq=1 ttl=64 time=270 ms
64 bytes from 172.24.63.248: icmp_seq=2 ttl=64 time=250 ms
64 bytes from 172.24.63.248: icmp_seq=3 ttl=64 time=277 ms
64 bytes from 172.24.63.248: icmp_seq=4 ttl=64 time=257 ms
64 bytes from 172.24.63.248: icmp_seq=5 ttl=64 time=260 ms
64 bytes from 172.24.63.248: icmp_seq=6 ttl=64 time=268 ms
64 bytes from 172.24.63.248: icmp_seq=7 ttl=64 time=284 ms
64 bytes from 172.24.63.248: icmp_seq=8 ttl=64 time=256 ms
^C
--- 172.24.63.248 ping statistics ---
9 packets transmitted, 8 received, 11.1111% packet loss, time 8302ms
rtt min/avg/max/mdev = 249.516/265.022/283.561/10.824 ms
pes1ug23cs375@Ubuntu22:~/Assignment-2$ python3 producer1.py
```

2.CONSUMER 1 :

```
pes1ug23cs196@pes1ug23cs196:~/big_data_asst$ python3 consumer1.py
✓ Consumer1 started - listening to 2 topics - cpu and mem.
Writing to consumer1_cpu_data.csv and consumer1_mem_data.csv files
```

3.CONSUMER 1 Spark job :

```

pes1ug23cs196@pes1ug23cs196:~/big_data_asst$ spark-submit spark_job1.py
25/10/22 05:39:42 WARN Utils: Your hostname, pes1ug23cs196 resolves to a loopback address: 127.0.1.1; using 192.168.64.7 instead (on interface enp0s1)
25/10/22 05:39:42 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
25/10/22 05:39:43 INFO SparkContext: Running Spark version 3.5.3
25/10/22 05:39:43 INFO SparkContext: OS info Linux, 5.15.0-157-generic, aarch64
25/10/22 05:39:43 INFO SparkContext: Java version 1.8.0_462
25/10/22 05:39:43 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
25/10/22 05:39:43 INFO ResourceUtils: =====
=====
25/10/22 05:39:43 INFO ResourceUtils: No custom resources configured for spark.driver.
25/10/22 05:39:43 INFO ResourceUtils: =====
=====
25/10/22 05:39:43 INFO SparkContext: Submitted application: SparkJob1_CPU_MEMORY_Anomaly
25/10/22 05:39:43 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(cores -> name: cores, amount: 1, script: , vendor: , memory -> name: memory, amount: 1024, script: , vendor: , offHeap -> name: offHeap, amount: 0, scri

25/10/22 05:45:16 INFO NettyBlockTransferService: Server created on 192.168.64.7:46177
25/10/22 05:45:16 INFO BlockManager: Using org.apache.spark.storage.RandomBlockReplicationPolicy for block replication policy
25/10/22 05:45:16 INFO BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 192.168.64.7, 46177, None)
25/10/22 05:45:16 INFO BlockManagerMasterEndpoint: Registering block manager 192.168.64.7:46177 with 366.3 MiB RAM, BlockManagerId(driver, 192.168.64.7, 46177, None)
25/10/22 05:45:16 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.64.7, 46177, None)
25/10/22 05:45:16 INFO BlockManager: Initialized BlockManager: BlockManagerId(driver, 192.168.64.7, 46177, None)
✓ Spark Session created for CPU and Memory anomaly detection.
✓ Spark Job 1 completed successfully.
Final Row Count: 14400
Output saved to: /home/pes1ug23cs196/big_data_asst/team_126_CPU_MEMORY.csv
pes1ug23cs196@pes1ug23cs196:~/big_data_asst$
```

4.CONSUMER 2 :

```

hadoop@harshaa-VirtualBox:~/assn2_2$ python3 consumer2.py
✓ Consumer2 started - listening to topic-net and topic-disk
Writing to consumer2_net_data.csv and consumer2_disk_data.csv
```

5.CONSUMER 2 Spark job :

```
hadoop@harshaa-VirtualBox:~/assn2_1$ spark-submit spark_job2.py
25/10/22 17:51:07 WARN Utils: Your hostname, harshaa-VirtualBox resolves to a loopback address: 127.0.1.1; using 10.0.2.15 instead (on interface enp0s3)
25/10/22 17:51:07 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
25/10/22 17:51:12 INFO SparkContext: Running Spark version 3.5.3
25/10/22 17:51:12 INFO SparkContext: OS info Linux, 6.14.0-33-generic, amd64
25/10/22 17:51:12 INFO SparkContext: Java version 1.8.0_462
25/10/22 17:51:13 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
25/10/22 17:51:13 INFO ResourceUtils: =====
25/10/22 17:51:13 INFO ResourceUtils: No custom resources configured for spark.driver.
25/10/22 17:51:13 INFO ResourceUtils: =====
25/10/22 17:51:13 INFO SparkContext: Submitted application: SparkJob2_NETDISK_Anomaly
25/10/22 17:51:13 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(cores -> name: cores, amount: 1, script: , vendor: , memory -> name: memory, amount: 10
24, script: , vendor: , offHeap -> name: offHeap, amount: 0, script: , vendor: ), task resources: Map(cpus -> name: cpus, amount: 1.0)
25/10/22 17:51:13 INFO ResourceProfile: Limiting resource is cpu
25/10/22 17:51:14 INFO ResourceProfileManager: Added ResourceProfile id: 0
25/10/22 17:51:14 INFO SecurityManager: Changing view acls to: hadoop
25/10/22 17:51:14 INFO SecurityManager: Changing modify acls to: hadoop
25/10/22 17:51:14 INFO SecurityManager: Changing view acls groups to:
25/10/22 17:51:14 INFO SecurityManager: Changing modify acls groups to:
25/10/22 17:51:14 INFO SecurityManager: authentication disabled; ui acls disabled; users with view permissions: hadoop; groups with view permissions: EMPTY; users with
modify permissions: hadoop; groups with modify permissions: EMPTY
25/10/22 17:51:15 INFO Utils: Successfully started service 'sparkDriver' on port 33559.
25/10/22 17:51:15 INFO SparkEnv: Registering MapOutputTracker
25/10/22 17:51:15 INFO SparkEnv: Registering BlockManagerMaster
25/10/22 17:51:15 INFO BlockManagerMasterEndpoint: Using org.apache.spark.storage.DefaultTopologyMapper for getting topology information
25/10/22 17:51:15 INFO BlockManagerMasterEndpoint: BlockManagerMasterEndpoint up
25/10/22 17:51:15 INFO SparkEnv: Registering BlockManagerMasterHeartbeat
25/10/22 17:51:16 INFO DiskBlockManager: Created local directory at /tmp/blockmgr-b83cb153-6c79-4a9f-b21d-bb3ab29a620d
25/10/22 17:51:16 INFO MemoryStore: MemoryStore started with capacity 366.3 MiB
25/10/22 17:51:16 INFO SparkEnv: Registering OutputCommitCoordinator
25/10/22 17:51:16 INFO JettyUtils: Start Jetty 0.0.0:4040 for SparkUI
25/10/22 17:51:16 INFO Utils: Successfully started service 'SparkUI' on port 4040.
25/10/22 17:51:17 INFO Executor: Starting executor ID driver on host 10.0.2.15
25/10/22 17:51:17 INFO Executor: OS info Linux, 6.14.0-33-generic, amd64
25/10/22 17:51:17 INFO Executor: Java version 1.8.0_462
25/10/22 17:51:17 INFO Executor: Starting executor with user classpath (userClassPathFirst = false): ''
```

```
25/10/22 17:51:13 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(cores -> name: cores, amount: 1, script: , vendor: , memory -> name: memory, amount: 10
24, script: , vendor: , offHeap -> name: offHeap, amount: 0, script: , vendor: ), task resources: Map(cpus -> name: cpus, amount: 1.0)
25/10/22 17:51:13 INFO ResourceProfileManager: Added ResourceProfile id: 0
25/10/22 17:51:14 INFO SecurityManager: Changing view acls to: hadoop
25/10/22 17:51:14 INFO SecurityManager: Changing modify acls to: hadoop
25/10/22 17:51:14 INFO SecurityManager: Changing view acls groups to:
25/10/22 17:51:14 INFO SecurityManager: Changing modify acls groups to:
25/10/22 17:51:14 INFO SecurityManager: authentication disabled; ui acls disabled; users with view permissions: hadoop; groups with view permissions: EMPTY; users with
modify permissions: hadoop; groups with modify permissions: EMPTY
25/10/22 17:51:15 INFO Utils: Successfully started service 'sparkDriver' on port 33559.
25/10/22 17:51:15 INFO SparkEnv: Registering MapOutputTracker
25/10/22 17:51:15 INFO SparkEnv: Registering BlockManagerMaster
25/10/22 17:51:15 INFO BlockManagerMasterEndpoint: Using org.apache.spark.storage.DefaultTopologyMapper for getting topology information
25/10/22 17:51:15 INFO BlockManagerMasterEndpoint: BlockManagerMasterEndpoint up
25/10/22 17:51:15 INFO SparkEnv: Registering BlockManagerMasterHeartbeat
25/10/22 17:51:16 INFO DiskBlockManager: Created local directory at /tmp/blockmgr-b83cb153-6c79-4a9f-b21d-bb3ab29a620d
25/10/22 17:51:16 INFO MemoryStore: MemoryStore started with capacity 366.3 MiB
25/10/22 17:51:16 INFO SparkEnv: Registering OutputCommitCoordinator
25/10/22 17:51:16 INFO JettyUtils: Start Jetty 0.0.0:4040 for SparkUI
25/10/22 17:51:17 INFO Utils: Successfully started service 'SparkUI' on port 4040.
25/10/22 17:51:17 INFO Executor: Starting executor ID driver on host 10.0.2.15
25/10/22 17:51:17 INFO Executor: OS info Linux, 6.14.0-33-generic, amd64
25/10/22 17:51:17 INFO Executor: Java version 1.8.0_462
25/10/22 17:51:17 INFO Executor: Starting executor with user classpath (userClassPathFirst = false): ''
25/10/22 17:51:17 INFO Executor: Created or updated repl class loader org.apache.spark.util.MutableURLClassLoader@5a567b5a for default.
25/10/22 17:51:17 INFO NettyBlockTransferService: Server created on 10.0.2.15:37055
25/10/22 17:51:17 INFO BlockManager: Using org.apache.spark.storage.RandomBlockReplicationPolicy for block replication policy
25/10/22 17:51:17 INFO BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 10.0.2.15, 37055, None)
25/10/22 17:51:17 INFO BlockManagerMasterEndpoint: Registering block manager 10.0.2.15:37055 with 366.3 MiB RAM, BlockManagerId(driver, 10.0.2.15, 37055, None)
25/10/22 17:51:17 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 10.0.2.15, 37055, None)
25/10/22 17:51:17 INFO BlockManager: Initialized BlockManager BlockManagerId(driver, 10.0.2.15, 37055, None)
✓ Spark Session created for Network and Disk anomaly detection.
✓ Spark Job 2 completed successfully.
Final Row Count: 14400
Output saved to: /home/hadoop/assn2_1/team_126_NET.DISK
hadoop@harshaa-VirtualBox:~/assn2_1$
```