



Roll No:

BTECH**(SEM. VI) THEORY EXAMINATION 2025-26****STREAM PROCESSING AND ANALYTICS****TIME: 3 HRS****M.MARKS: 70****Note:** Attempt all Sections. In case of any missing data; choose suitably.**SECTION A****1. Attempt all questions in brief.****02 x 7 = 14**

Q no.	Question	CO	Level
a.	Define Stream Processing and provide two real-world examples.	1	K1
b.	Differentiate between Stateless and Stateful processing.	1	K2
c.	What does latency represent in a stream processing engine?	2	K1
d.	Briefly explain the concept of micro-batching in Apache Spark.	3	K2
e.	Define event-time and how it differs from processing-time.	4	K1
f.	List any two default sinks available in structured streaming.	4	K1
g.	What is the purpose of a Spark Streaming Context (SSC)?	5	K2

SECTION B**2. Attempt any three of the following:****07 x 3 = 21**

Q no.	Question	CO	Level
a.	Discuss the concept of scaling up vs. scaling out in the context of distributed stream processing.	1	K2
b.	Describe Spark's memory usage and its impact on latency and throughput.	2	K3
c.	Explain the Resilience model in Spark. How do RDDs guarantee fault tolerance?	3	K2
d.	Discuss the structured streaming programming model and the "Unbounded Table" concept.	4	K5
e.	Explain the Spark streaming execution model and the role of DStreams in materializing results.	5	K5

SECTION C**3. Attempt any one part of the following:****07 x 1 = 07**

Q no.	Question	CO	Level
a.	Describe the steps to install and configure Apache Spark with JDK and Scala IDE.	1	K2
b.	Explain the use of window aggregations and the effect of time on transformations	1	K2

4. Attempt any one part of the following:**07 x 1 = 07**

Q no.	Question	CO	Level
a.	Explain the architectural models of a data platform. How does a batch-processing component complement a streaming application?	2	K2
b.	Design a Spark streaming program to count words in text data received from a TCP socket.	2	K3



Paper ID : 261566

Printed Page: 2 of 2

Subject Code:BCAM062

Roll No:

BTECH

(SEM. VI) THEORY EXAMINATION 2025-26

STREAM PROCESSING AND ANALYTICS

TIME: 3 HRS

M.MARKS: 70

5. Attempt any *one* part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	What are Resilient Distributed Datasets (RDDs)? Explain how the lineage graph helps in distributed processing.	3	K2
b.	Analyze the micro-batching processing model. How does it simplify fault tolerance compared to continuous processing?	3	K4

6. Attempt any *one* part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Explain the structured streaming programming model. How does it treat a live data stream as a continuously appended table?	4	K2
b.	Compare micro-batch processing with continuous processing modes in structured streaming. When would you prefer one over the other?	4	K4

7. Attempt any *one* part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Explain the concept of time-based stream processing in DStreams. How do slide intervals and window intervals interact?	5	K2
b.	Discuss performance tuning strategies for Spark streaming. How does adjusting the batch interval affect latency and throughput?	5	K4