UG

RV College of Engineering®
(Autonomous Institution affiliated to VTU)
Bengaluru-560 059
Course Title : Stream Processing and Analytics
Course Code : AI254TA

Model Question Paper

COURSE CODE: **21AI72**	**SEM**: VII
COURSE TITLE : **Stream Processing and Analytics**	
Duration of Paper: 03 Hrs	

Instructions to Candidates:
1. Answer all questions from Part A
2. Any 5 Full questions from Part B choosing one from each side. (Question No.2 is compulsory)

Question No	PART A	Marks	BTL	CO
1.1	Classify the soft and near real time systems with an example	2	1	1
1.2	What is check pointing ?	2	1	1
1.3	Give any two needs for message queueing tier?	2	2	1
1.4	Draw the stream processing framework of apache storm?	2	1	3
1.5	Give the difference between stream time vs event time	02	2	2
1.6	Define a batch loader. Give any one benefit of the same	2	2	2
1.7	Give any two drawbacks of data sync	2	2	2
1.8	What is the role of "Mirror Maker" in kafka?	2	2	3
1.9	"There are many choices for publish/subscribe messaging systems ". Examine the statement and give any two reasons to support Kafka in this context	2	4	3
1.10	Differentiate Internal Vs External State in stream processing ?	2	2	2
	PART B			
2(a)	With a neat diagram, elaborate the streaming data architecture	**10**	**02**	**01**
2(b)	Give the importance of scaling in detail	**06**	**02**	**01**
	Unit II			
3(a)	Elaborate the need for message queueing tier. Give relevant example	**10**	**02**	**01**
3(b)	Imagine you have built a real-time traffic-routing system that allows people driving around any city to use your smartphone app to get updates and be re-routed based on up-to-the-moment traffic conditions. Three months pass, and now your business wants to offer a historical traffic-replay product that lets a user pick a city and replay the traffic data for a given day,	**06**	**03**	**03**

	week, or month. For the given scenario, build a message queueing model and explain the steps in detail			
	OR			
4(a)	Elaborate the seven points of failure in stream processing data flow with a real time example	**10**	**02**	**01**
4(b)	For the scenario given in 3b, identify and apply the message delivery semantics. With a neat diagram explain the concept implemented	06	03	02
	Unit III			
5(a)	Imagine we are collecting data from a fitness-tracking device such as a Fitbit, and the data is flowing into our streaming system. Stream time would be when the fitness event enters the analysis tier; event time would be when it takes place on the device. The usage head of product marketing has asked us to build a dashboard that shows the average speed for all runners broken down by age groups, such as 12–17, 18–24, 25–34, and so on. The dashboard should be updated every 5 seconds, and the averages should represent data for the last 30 minutes. Identify and elaborate the type of window technique that can be used for the same. Justify your answer.	10	04	03
5(b)	Consider the flow of data in an digital marketing system the server receives millions of views per minute and there is a need for statistical analysis of data. How do we take a random sample from a data set that you can't hold in memory or on disk and perform analysis?	06	03	03
	OR			
6(a)	With a relevant example, discuss the different caching mechanism for stream data services.	10	03	02
6(b)	How do different stream processing tools handle sliding window protocol? Explain with an example	06	03	03
	Unit IV			
7(a)	Elaborate any two Protocols used to send data to the client. Support the answer with relevant examples	08	02	03
7(b)	There are many choices for publish/subscribe messaging systems, so what makes Apache Kafka a good choice?	08	02	03
	OR			
8(a)	What are the different potential ways to think about integrating the different types of filtering for each of the protocols	08	02	03
8(b)	Elucidate the data ecosystem of kafka with a neat diagram	08	02	03
	Unit V			
9(a)	Consider a Real-Time Fraud Detection in Online Transactions. If a data pipeline is to be built using kafka, list and elaborate the different considerations during the implementation .	10	03	03
9(b)	How does kafka connect deal with Connectors	08	02	03
	OR			
10(a)	consider an e-commerce website that generates order-placed events. When a customer places an order, an event is produced with details like order ID, customer	10	03	03

	ID, product details, and total amount. The goal is to process each order event independently and send an order confirmation email to the customer. Apply Single event processing design pattern and explain the steps in detail			
10(b)	How does kafka handle the different types of joints? Give examples	08	03	03

Name of the Scrutinizer

Signature of Scrutinizer

Name of the BoE Chairperson

Name of the BoE Chairperson