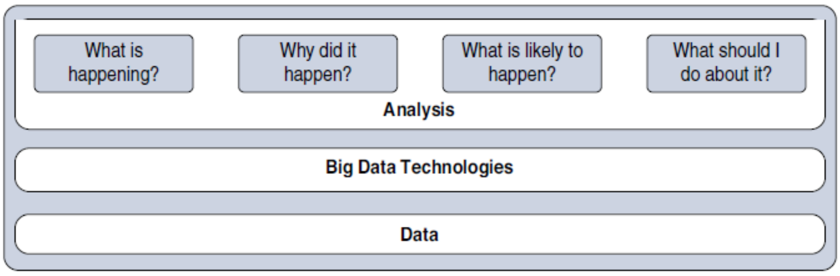
USN | | | | | | | | | | |

Internal Assessment Test 3 – MAY 2023

Sub:	Internet of Things Technology				Sub Code:	18CS81	Branch:	CSE
Date:	13/5/23	Duration:	1.5 hrs	Max Marks: 50	Sem / Sec:	VIII Sem A/B/C		OBE

SCHEME & SOLUTIONS

		MARKS	CO	RBT
1	Explain different types of data analysis results wrt the diagram.	[10]	CO4	L3



Explain 4 types – 2.5 * 4

Ans

Descriptive: Descriptive data analysis tells you what is happening, either now or in the past. For example, a thermometer in a truck engine reports temperature values every second. From a descriptive analysis perspective, you can pull this data at any moment to gain insight into the current operating condition of the truck engine. If the temperature value is too high, then there may be a cooling problem or the engine may be experiencing too much load.

Diagnostic: When you are interested in the “why,” diagnostic data analysis can provide the answer. Continuing with the example of the temperature sensor in the truck engine, you might wonder why the truck engine failed. Diagnostic analysis might show that the temperature of the engine was too high, and the engine overheated. Applying diagnostic analysis across the data generated by a wide range of smart objects can provide a clear picture of why a problem or an event occurred.

Predictive: Predictive analysis aims to foretell problems or issues before they occur. For example, with historical values of temperatures for the truck engine, predictive analysis could provide an estimate on the remaining life of certain components in the engine. These components could then be proactively replaced before failure occurs. Or perhaps if temperature values of the truck engine start to rise slowly over time, this could indicate the need for an oil change or some other sort of engine cooling maintenance.

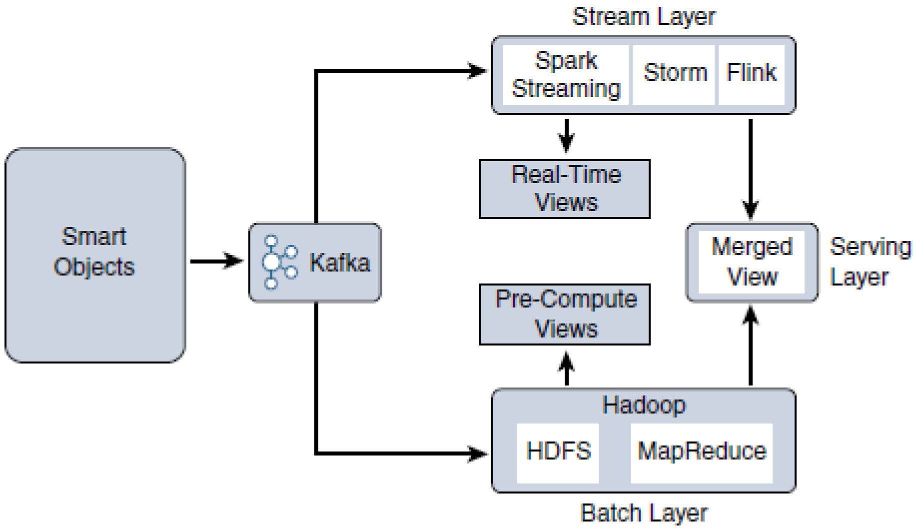
Prescriptive: Prescriptive analysis goes a step beyond predictive and recommends solutions for upcoming problems. A prescriptive analysis of the temperature data from a truck engine might calculate various alternatives to cost-effectively maintain our truck. These calculations could range from the cost necessary for more frequent oil changes and cooling maintenance to installing new cooling equipment on the engine or upgrading to a lease on a model with a more powerful engine. Prescriptive analysis looks at a variety of factors and makes the appropriate recommendation.

		MARKS	CO	RBT
2	What is Apache spark? Explain layers in lambda architecture. Explain Spark - 5 marks Lamda architecture with diagram – 5 marks	[10]	CO4	L2

Ans

Apache Spark is an in-memory distributed data analytics platform designed to accelerate processes in the Hadoop ecosystem. The “in-memory” characteristic of Spark is what enables it to run jobs very quickly. At each stage of a MapReduce operation, the data is read and written back to the disk, which means latency is introduced through each disk operation. However, with Spark, the processing of this data is moved into high-speed memory, which has significantly lower latency. This speeds the batch processing jobs and also allows for near-real-time processing of events.

Real-time processing is done by a component of the Apache Spark project called Spark Streaming. Spark Streaming is an extension of Spark Core that is responsible for taking live streamed data from a messaging system, like Kafka, and dividing it into smaller microbatches. These microbatches are called discretized streams, or DStreams. The Spark processing engine is able to operate on these smaller pieces of data, allowing rapid insights into the data and subsequent actions. Due to this "instant feedback" capability, Spark is becoming an important component in many IoT deployments. Systems that control safety and security of personnel, time-sensitive processes in the manufacturing space, and infrastructure control in traffic management all benefit from these real-time streaming capabilities.



The key elements of a data infrastructure to support many IoT use cases involve the collection, processing, and storage of data using multiple technologies. Querying both data in motion (streaming) and data at rest (batch processing) requires a combination of the Hadoop ecosystem projects discussed. One architecture that is currently being leveraged for this functionality is the Lambda Architecture.

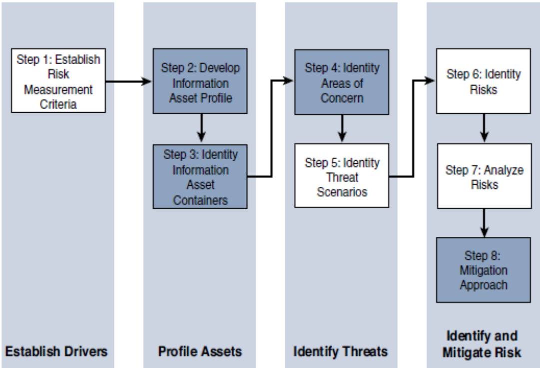
Lambda is a data management system that consists of two layers for ingesting data (Batch and Stream) and one layer for providing the combined data (Serving). These layers allow for the packages discussed previously, like Spark and MapReduce, to operate on the data independently, focusing on the key attributes for which they are designed and optimized. Data is taken from a message broker, commonly Kafka, and processed by each layer in parallel, and the resulting data is delivered to a data store where additional processing or queries can be run.

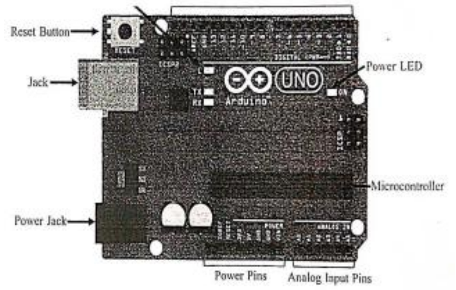
■ **Stream layer**: This layer is responsible for near-real-time processing of events. Technologies such as Spark Streaming, Storm, or Flink are used to quickly ingest, process, and analyze data on this layer. Alerting and automated actions can be triggered on events that require rapid response or could result in catastrophic outcomes if not handled immediately.

■ **Batch layer**: The Batch layer consists of a batch-processing engine and data store.

If an organization is using other parts of the Hadoop ecosystem for the other layers, MapReduce and HDFS can easily fit the bill. Other database technologies, such as MPPs, NoSQL, or data warehouses can also provide what is needed by this layer.

■ **Serving layer**: The Serving layer is a data store and mediator that decides which

	of the ingest layers to query based on the expected result or view into the data. If an aggregate or historical view is requested, it may invoke the Batch layer. If real-time analytics is needed, it may invoke the Stream layer. The Serving layer is often used by the data consumers to access both layers simultaneously.			
3	Explain OCTAVE Allegro steps and phases with a neat diagram Steps & Phases – 7 marks Diagram – 3 marks	[10]	CO4	L2
Ans	OCTAVE (Operationally Critical Threat, Asset and Vulnerability Evaluation) is a from risk assessment frameworks the Software Engineering Institute at Carnegie Mellon University.  The **first step** of the OCTAVE Allegro methodology is to establish a risk measurement criterion. OCTAVE provides a fairly simple means of doing this with an emphasis on impact, value, and measurement. The point of having a risk measurement criterion is that at any point in the later stages, prioritization can take place against the reference model. The **second step** is to develop an information asset profile. This profile is populated with assets, a prioritization of assets, attributes associated with each asset, including owners, custodians, people, explicit security requirements, and technology assets. It is important to stress the importance of process. The **third step** is to identify information asset containers. Roughly speaking, this is the range of transports and possible locations where the information might reside. This references the compute elements and the networks by which they communicate. However, it can also mean physical manifestations such as hard copy documents or even the people who know the information. The **fourth step** is to identify areas of concern. At this point, we depart from a data flow, touch, and attribute focus to one where judgments are made through a mapping of security-related attributes to more business-focused use cases. At this stage, the analyst looks to risk profiles and delves into the previously mentioned risk analysis. It is no longer just facts, but there is also an element of creativity that can factor into the evaluation. The **fifth step** is where threat scenarios are identified. Threats are broadly (and properly) identified as potential undesirable events. This definition means that results from both malevolent and accidental causes are viable threats. In the context of operational focus, this is a valuable consideration. It is at this point that an explicit identification of actors, motives, and outcomes occurs. At the **sixth step** risks are identified. Within OCTAVE, risk is the possibility of an undesired outcome. This is extended to focus on how the organization is impacted. For more focused analysis, this can be localized, but the potential impact to the organization could extend outside the boundaries of the operation. The **seventh step** is risk analysis, with the effort placed on qualitative evaluation of the impacts of the risk. Here the risk measurement criteria defined in the first step are explicitly brought into the process.			

4	Explain the different pins and parts of Arduino Uno board. Diagram – 2 marks Explanation of pins & ports – 8 marks	[10]	CO5	L2
Ans	 **5v and 3.3v** They provide regulated 5 and 3.3v to power external components according to manufacturer specifications. **GND** In the Arduino Uno pinout, you can find 5 GND pins, which are all interconnected. The GND pins are used to close the electrical circuit and provide a common logic reference level throughout your circuit. **RESET** - resets the Arduino **IOREF** - This pin is the input/output reference. It provides the voltage reference with which the microcontroller operates. **Arduino Pins A0-A5** are capable of reading analog voltages. **Pins 0-13** of the Arduino Uno serve as digital input/output pins. **Pin 13** of the Arduino Uno is connected to the built-in LED. In the Arduino Uno - pins 3,5,6,9,10,11 have **PWM capability**. **USB Connection** – Used for powering up your Arduino and uploading sketches. **TX/RX** – Transmit and receive data indication LEDs. **ATmega Microcontroller** – This is the brains and is where the programs are stored. **Voltage Regulator** – This controls the amount of voltage going into the Arduino board. **DC Power Barrel Jack** – This is used for powering your Arduino with a power supply			
5	Explain the basic structure of Arduino programming. Write a program to blink LED for Arduino Uno. Structure – 5 marks Program – 5 makrs	[10]	CO5	L3

Ans	Sketch − The first new terminology is the Arduino program called “sketch”.
	Arduino programs can be divided in three main parts: Structure, Values (variables and constants), and Functions.
	In software, two required functions / methods / routines for Arduino programming are:
	
	void setup()
	{
	    // runs once
	}
	
	void loop()
	{
	    // repeats
	}			

The setup() function is called when a sketch starts. Use it to initialize the variables, pin modes, start using libraries, etc. The setup function will only run once, after each power up or reset of the Arduino board.

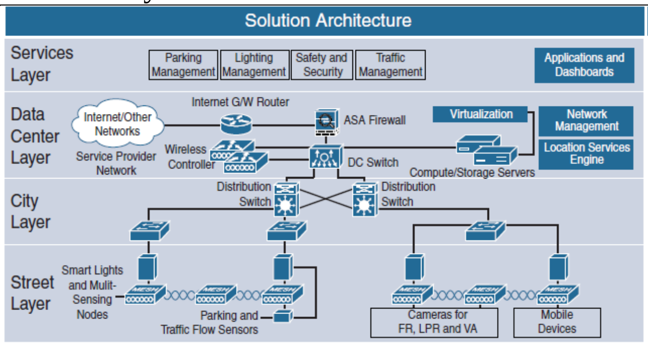
After creating a setup() function, which initializes and sets the initial values, the loop() function does precisely what its name suggests, and loops consecutively, allowing your program to change and respond. Use it to actively control the Arduino board.

```
void setup() {
        // set Pin 3 to output
        pinMode(3, OUTPUT);
}

void loop() {
        digitalWrite(3, HIGH);  // turn LED on (output 5V)
        delay(1000);            // wait one second
        digitalWrite(3, LOW);   // turn LED off (output 0V)
        delay(1000);            // wait another second
}
```

6	With a case study explain key verticals targeted in smart cities. Discuss all layers – 2.5 * 4	[10]	CO5	L3

Ans



Street Layer

The street layer is composed of devices and sensors that collect data and take action based on instructions from the overall solution, as well as the networking components needed to aggregate and collect data.

City Layer

At the city layer, which is above the street layer, network routers and switches must be deployed to match the size of city data that needs to be transported. This layer aggregates all data collected by sensors and the end-node network into a single transport network.

Data Center Layer

Ultimately, data collected from the sensors is sent to a data center, where it can be processed and correlated. Based on this processing of data, meaningful information and trends can be derived, and information can be provided back. For example, an application in a data center can provide a global view of the city traffic and help authorities decide on the need for more or less common transport vehicles. At the same time, an automated response can be generated. For example, the same traffic information can be processed to automatically regulate and coordinate the street light durations at the scale of the entire city to limit traffic congestion.

Services Layer

Ultimately, the true value of ICT connectivity comes from the services that the measured data can provide to different users operating within a city. Smart city applications can provide value to and visibility for a variety of user types, including

	city operators, citizens, and law enforcement. The collected data should be visualized according to the specific needs of each consumer of that data and the particular user experience requirements and individual use cases.			