

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI (RAJ.)**  
**I SEMESTER 2019-2020**  
**TERM PROJECT**

**Course No.: CS/SS G527**  
**Weight: 60M (15%)**

**Course Title: Cloud Computing**

**Note:**

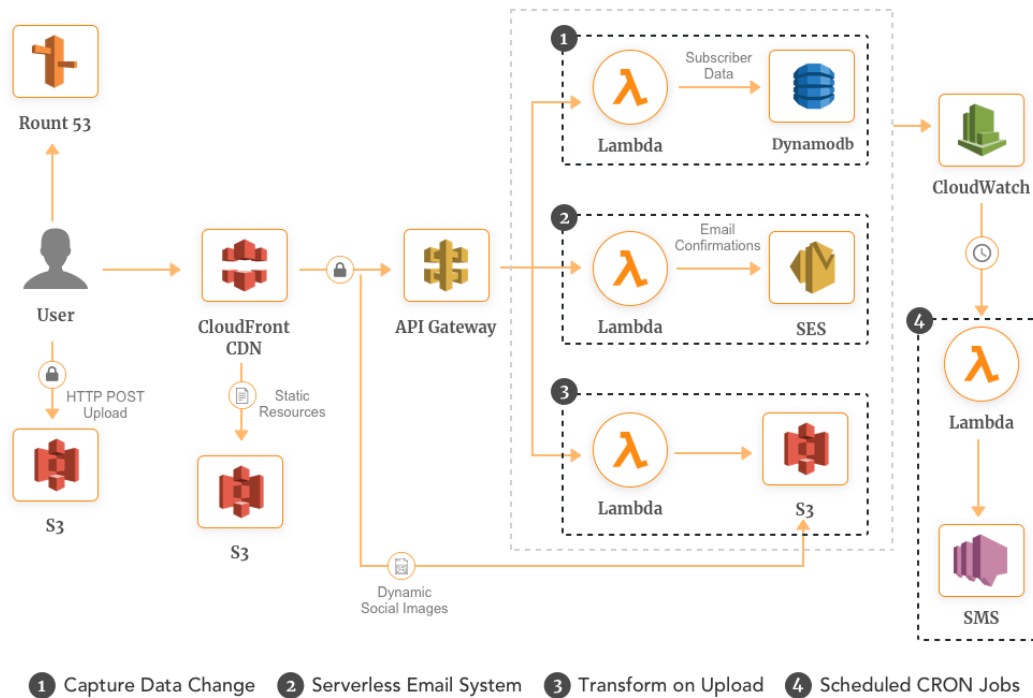
- Should be done in a group of maximum four students
- Project has two phases (i) System Design (ii) System Implementation
- Choose your demo slot here:

[Google Sheet](#)

**Design and Implementation of Platform as a service for Pipelined Execution**

Goal of the Term Project is to build a platform-as-a-service for enabling pipelined execution. In [pipelined execution](#), the output of one service is input to the next service. Platform should offer a variety of services, and the user chooses the services and also the order in which they have to be applied to the data. For example, in AWS, the incoming data can be made to pass through services such as Lambda, Kinesis, DynamoDB, Redshift, SNS, SQS etc in the selected order. The output of one service can be made as input to multiple services.





Term project will have the following phases:

### 1. System Design

- In this phase, as first step, identify the domain of the problem, what kind of data stream is expected by your platform, and what services your platform will offer are to be identified.
- For each identified service,
  - specify whether it should be monolithic or a set of micro services.
  - specify its needs such as storage, database etc
  - specify the format of input it takes
  - explain the coarse-grained steps it will follow
  - specify the format of output it gives
- Identify what technical alternatives exist for connecting the services in the pipeline and chose one with justification.
- Auto Scaling:
  - Compute your system's capacity.
  - Identify the parameters for deciding on auto-scaling
  - Explain how do you monitor these parameters
  - Identify ways to optimize resources with auto-scale.
- Infrastructure:
  - Will your services run in VMs or in containers? Justify.
  - When multiple users identify different pipelines with streams at different data rates, it becomes very important to optimize your resource usage. Explain resource monitoring and resource scheduling of your platform.
- SLA

- SLA parameters
  - SLA monitoring
- Availability
  - Compute your system's availability
  - How do you deal with failures?
- Pricing
  - Design a pricing mechanism for your service?
  - How will you monitor?

Deliverables: Design document containing answers to all the above stated aspects including other details which are not covered above.

Submission Deadline: 13th October

Submission Mode: To be uploaded on <http://nalanda.bits-pilani.ac.in>

## 2. System Implementation

- Design stated in phase1 should be Implemented. Identify the programming platform for the implementation.
- The platform-as-a-service should be demonstrated on a cluster of at least 3 physical machines.
- Your platform should provide two interfaces: (i) an interface to the user to create a pipeline (ii) an interface for the cloud admin for pricing, listing resources, current instances, current users, SLA monitoring etc
- Your platform should also show/log
  - the intermediate input/output and execution details for each user stream
  - number of service instances being run for each service for each user stream
  - response time for each data item
- The data stream can be generated using *netcat* and *pv* tools. The data rate can be increased dynamically to check the system response.

Deliverables: Source code, PDF document explaining implementation details, README file on how to execute

Submission deadline: 17<sup>th</sup> November

Submission Mode: Upload through <http://nalanda.bits-pilani.ac.in> in a single ZIP file