Page No.
Date : / /

Workflow Engine For Clouds

* Workflow Engine in Cloud:

=> This workflow engine coordinates tasks and schedules them on remote resource.

This are the main part of the workflow engine.

1 Workflow management System:

The main function of WFMS is to schedules jobs in the workflow to remote resource based on the user's specified Quality of Service.

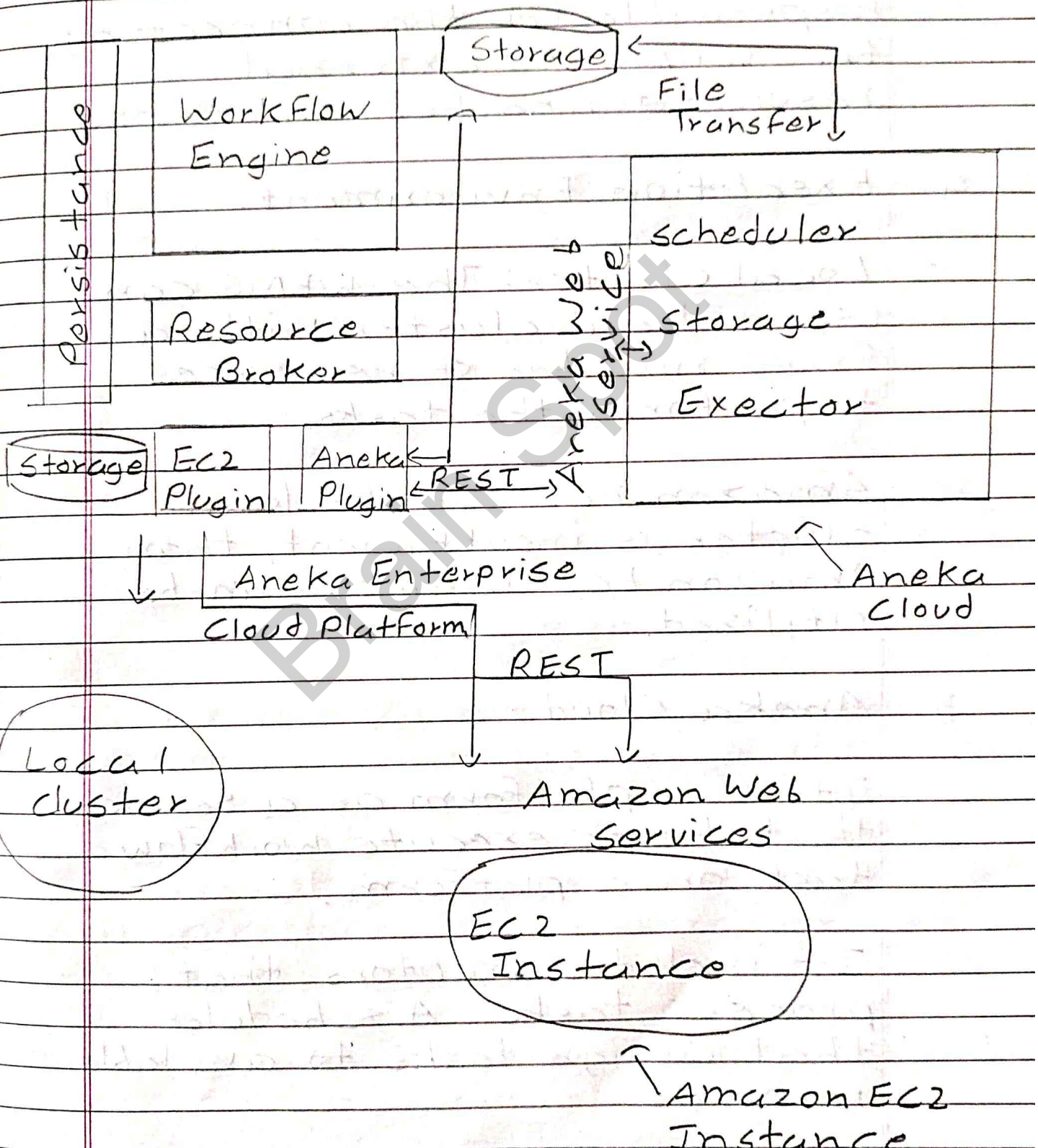
It contains serval components such as Resource Broker and plug-ins for various platforms like Amazon EC2 and Aneka Cloud.

2 Storage Service:

The storage service temporaliy stores component of the application.

Brain Spot

Storage can be provided by Amazon S3 or FTP & FTP servers.



Brain Spot

Date : / /

3 Resource Broker:

The Resource Broker is responsible for the connecting the WFMS to external resource like EC2 and Aneka

4 Execution Environment:

- Local Cluster: The WFMS can use a local cluster with a fixed number of resources to perform the tasks.

- Amazon EC2: If the local cluster is insufficient, than Amazon EC2 intanse can be utilized.

5 Aneka Cloud:

It is a platform as a Service that helps execute workflow task on a platform.

It includes excutors that process tasks, A scheduler that assign tasks to available resource.

Brain Spot

6 File Transfer:

File Transfer mechansim used to transfer WFMS data between the storage service and the Aneka cloud and EC2 instances.

7 Web Services and REST APIs:

- Aneka Web Service: This services provides API for job submission, job monitoring and resource negotiation.

- REST Interface: The system uses REST interface to communicate with various components.

* Workflow management System:

=> The architecture of Workflow management system (WFMS) is scientific application is designed to manage and execute complex workflow.

-> Main Part:

1 User Interface (Top Layer):

Brain Spot

SMVS
Page No.
Date : / /

This top layer interacts with the end users.

It includes tools for workflow composition, planning, submission and monitoring.

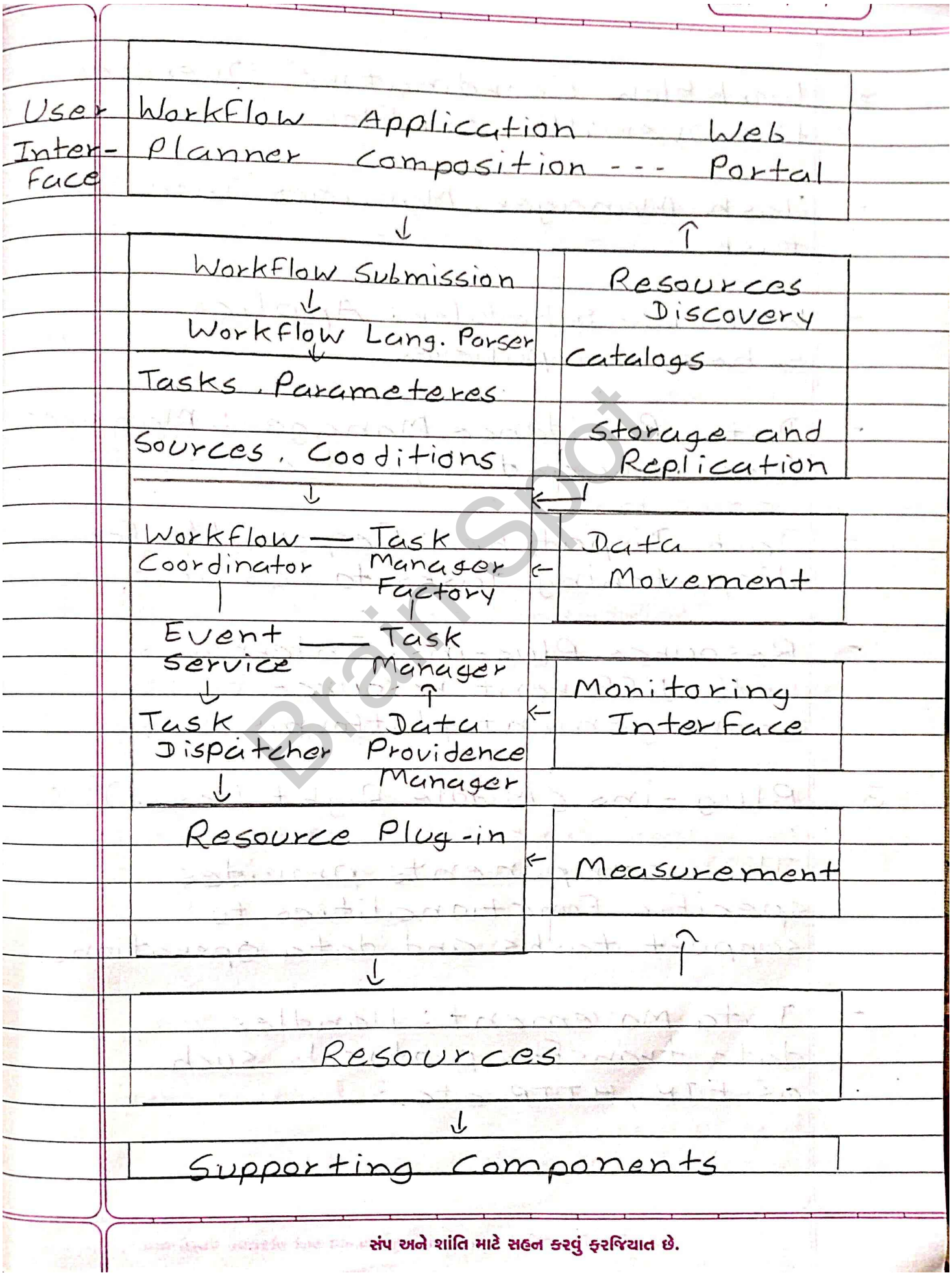
- Workflow Planner: Responsible for defining workflow.
- Application Composition: Helps in organizing and composing the workflow.
- Workflow Description: Handles the specific of tasks.

2 Engine (Middle Layer):

The core of the system where actual workflow management happens.

- Workflow Submission Handler: Handles the process of submitting the workflow to the system.
- Workflow Language Parser: Converts the workflow formates.

Brain Spot


User Interface
Workflow Planner
Application Composition
Web Portal
Workflow Submission
Workflow Lang. Parser
Tasks, Parameteres
Sources, Cooditions
Resources Discovery
Catalogs
Storage and Replication
Workflow Coordinator
Task Manager Factory
Data Movement
Event Service
Task Manager
Task Dispatcher
Data Providence Manager
Monitoring Interface
Resource Plug-in
Measurement
Resources
Supporting Components

Brain Spot

SMVS
Page No.
Date : / /

- Workflow Coordinators: Oversees the overall workflow.
- Task Manager: Manages Individua task.
- Workflow Scheduler: Applies scheduling policies.
- Data Providence Manager: Manage data flow and dependencies.
- Task Dispatcher: Responsible for dispatching tasks to resource.
- Resource Plug-ins: Interfaces with different resource management platforms.

3 Plug-ins (Middle Right Layer):

This components provides specific functionalities to support tasks and data operations.

- Data Movement: Handles data transfer protocals such as FTP, HTTP etc.

BrainSpot

- Monitoring Interface: Provides tools for real-time monitoring.
- Measurements: Trackes energy consumption, resource utilization and other metrics.

4 Resources (Bottom Layer):

At the lowest layer, there are resources like clusters, global grids that perform the actual computation and storage tasks.

5 Supporting Components:

- Storage and Replication: Manages data replication across different systems.
- Resource Utilization: Keeps track of resource and their availability.

Brain Spot

Page No.
Date : / /

* Evolutionary Multiobjective Optimizations (EMO):

=> EMO is a technique based on genetic algorithms used to optimize multiple search objectives.

EMO tasks are executed in parallel which reducing the total time required for the computation.

-> Objectives:

1 Design an Execution Model: Create a workflow that can utilze standard computing resource.

2 Parallel Tasks Execution: Split the task across multiple resource to reduce overall completion time.

3 Dynamic Resource Provisioning: Automatically allocate more resources when task increase in number.

Brain Spot

4 Repetitive Experiments: Allows the workflow to repeatedly run similar experiments.

5 Fault Tolerance and Result Managment: Handles execution failures and manage the workflow.

-> Workflow Solution:

1 Workflow structure:

The EMO workflow consists of several tasks organized in a branching structure.

Five different topology files are used for the iteration process.

2 Branching and Merging:

Each topology starts from root level with tasks.

As the task progress, the population from different ~~totog~~ topologies is merged.

Brain Spot

SMVS
Page No.
Date : / /

3 Task Levels:

The structure has 12 Levels of tasks.

After multiple iterations, the branches merge and split again at higher level until the final tasks is completed.

4 Parallelization:

The Branching structure allows for parallel execution of tasks on different topologies.

Brain Spot