AWS Batch Universal Job Scheduler

1. Job schedule definition (json document) is loaded into master table (one-time)
2. User can initiate a job schedule by requesting job via REST API (passing job id)
3. Optionally, a job schedule can be initiated by a cloud watch event at a specific time
4. State machine gets schedule, submits jobs to AWS Batch and monitors job
5. AWS Batch scheduling engine handles job dependencies and orchestration

6. Big data ETL jobs are executed in the data lake by AWS Glue, lambda, etc.
7. Individual EC2-based jobs can be executed via SSM agent
8. AWS batch container jobs can perform long-running transformations and other work
9. Job failure feedback is delivered via SNS
10. Job result (success/failure) is updated into transactions database for reporting

AWS

Job Schedule Definition

1

Analytics

Amazon QuickSight

bucket

TTL export

Job Request Transactions

Amazon DynamoDB

bucket

Job Schedule Definitions

Amazon DynamoDB

Schedule

Job final status

10

Cloudwatch Event

3

Schedule ID

Amazon SNS

Job Fail Notifications

9

4

AWS Step Functions

Schedule ID

Amazon API Gateway

Schedule ID

Users

2

Jobs

5

AWS Batch

ECR registry

Various custom job runner images

6 Big data jobs

7 EC2 jobs

8 Container jobs

Container jobs

Data Lake ETL

Remote Command

EC2

Data Lake

AWS Glue

Lambda

bucket

bucket

Processes