


Big Data on Cloud

Challenges with On Prem Solutions

- ↳ Scalability
- ↳ Maintenance
- ↳ Security
- ↳ Cost

Main Cloud Platforms

- ↳ AWS
- ↳ Azure
- ↳ GCP

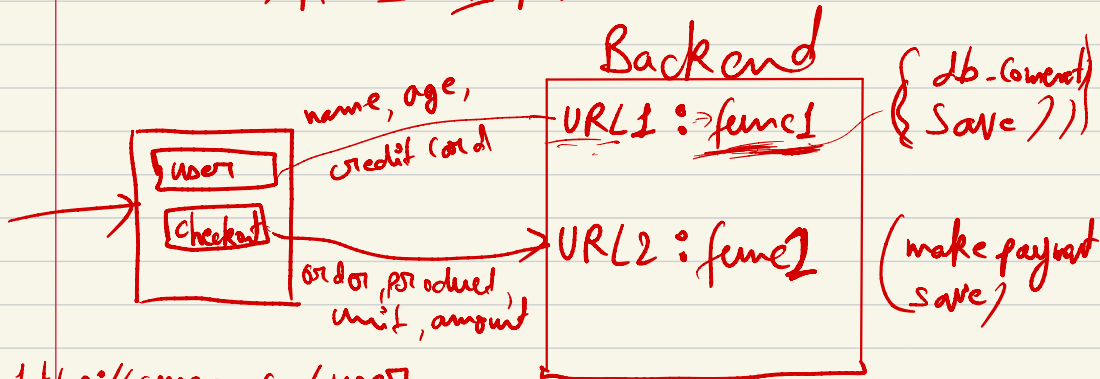
Other players

- IBM
- Oracle
- Alibaba

Important AWS Services for Big Data / Data Engineer

- ↳ AWS S3 (Simple Storage Service)
- ↳ Database Services
 - ↳ AWS RDS (Relational Database Service)
(MySQL, PostgreSQL)
 - ↳ NoSQL Databases
 - ↳ AWS DynamoDB (Key-Value)
- ↳ Serverless Service
 - ↳ AWS Lambda
- ↳ Serverless DB Service
 - ↳ AWS Athena
- ↳ AWS EMR (Managed cluster)
- ↳ AWS Glue (Managed ETL Service)
- ↳ For Infra
 - ↳ AWS EC2
- ↳ Data Warehouse
 - ↳ AWS Redshift

- ↳ for Scheduling / Monitoring
 - ↳ AWS CloudWatch
 - ↳ AWS EventBridge
- ↳ SNS (Simple Notification Service)
- ↳ SQS (Simple Queue Service)
- ↳ for Dashboarding
 - ↳ AWS QuickSight
- ↳ for authentication
 - ↳ AWS IAM



URL1 <http://amazon.com/user>

URL2 <http://amazon.com/checkout>

