

# Architectural Design

App	Recommender System Data Processing and Model Training	Data Processing and Model Training	Data Processing and Model Training Cost	Components and Services EMR	EMR Configuration	EMR Cost	Storage	Storage Cost	Data Transfer	Data Transfer Cost	Monitoring & Logging	Monitoring & Logging Cost
Web Mobile	Apache Spark with Spark Mllib. Employ collaborative filtering or content-based filtering algorithms. Real-time or batch processing to provide recommendations. Integration with the web/app backend to serve recommendations to users.	Model Training and Inference: Using EC2 instances (p2.xlarge for GPU). Training: 10 hours/month. Cost: \$0.90/hour * 10 = \$9/month	\$ 9.00	Apache Spark on Amazon EMR (Elastic MapReduce):  Cluster Configuration: Master Node: Manages the Spark cluster. Worker Nodes: Perform data processing and model training. Instance Types: Choose instances based on processing requirements. Master Node: m5.large (2 vCPUs. 8 GB RAM) Worker Nodes: m5.xlarge (4 vCPUs. 16 GB RAM)	Cluster Configuration:  Master Node: m5.large instance: \$0.096 per hour Worker Nodes: m5.xlarge instances: \$0.192 per hour Assume 4 worker nodes for sufficient processing power. EMR Costs:  Master Node: \$0.096/hour Monthly cost: \$0.096 * 24 * 30 = \$69.12 Worker Nodes: 4 * \$0.192/hour Monthly cost: 4 * \$0.192 * 24 * 30 = \$552.96	\$ 622.08	Storage for data and model artifacts. Assume 1 TB of storage. Cost: \$0.023 per GB-month Monthly cost: 1024 GB * \$0.023 = \$23.55	\$ 23.55	Assume 1 TB of data transfer per month. First 1 GB/month is free. Additional data transfer: \$0.09 per GB Monthly cost: (1024 - 1) * \$0.09 = \$92.07	\$ 92.07	AWS CloudWatch:  Basic monitoring is free. Additional charges for detailed monitoring and logs. Assume \$20/month for detailed monitoring and logs.	\$ 20.00