

# EKS Cluster Deployment Architecture

## Custom Nodepool (role=custom-workload)

MCP Server  
3 replicas  
HPA

Chat Simulator  
1 replica  
LoadBalancer

Lightweight Filter  
2 replicas  
ClusterIP

Decision Handler  
2 replicas  
ClusterIP

## Default Nodepool

PostgreSQL  
StatefulSet

Redis  
Deployment

Prometheus  
Deployment

Grafana  
Deployment

DeepSeek LLM  
Existing

## Persistent Storage

PostgreSQL PVC  
10Gi

Prometheus PVC  
20Gi

Grafana PVC  
5Gi