

Deployment configuration

deployment_config:

Dockerfile configuration

dockerfile_path: "./Dockerfile" # The path to the Dockerfile for building the agent's image

dockerfile_port: 8080 # The port number the agent will listen on

Resource allocation for the agent

resources:

cpu: 2 # Number of CPUs allocated to the agent

memory: "2Gi" # Memory allocation for the agent in gigabytes

max_instances: 5 # Maximum number of instances to scale up to

min_instances: 1 # Minimum number of instances to keep running

timeout: 300s # Request timeout setting in seconds

Autoscaling configuration

autoscaling:

max_concurrency: 80 # Maximum number of requests the agent can handle concurrently

target_utilization: 0.6 # CPU utilization target for auto-scaling

Environment variables for the agent

environment_variables:

DATABASE_URL: "postgres://user:password@db-url" # URL for the database connection

API_KEY: "your-secret-api-key" # API key for authentication

LOG_LEVEL: "info" # Log level for the agent

Secrets configuration

secrets:

SECRET_NAME_1: "projects/my-project/secrets/my-secret/versions/latest" # Path to a secret