www.griddynamics.com

Grid Dynamics
trusted engineering partner for digital transformation

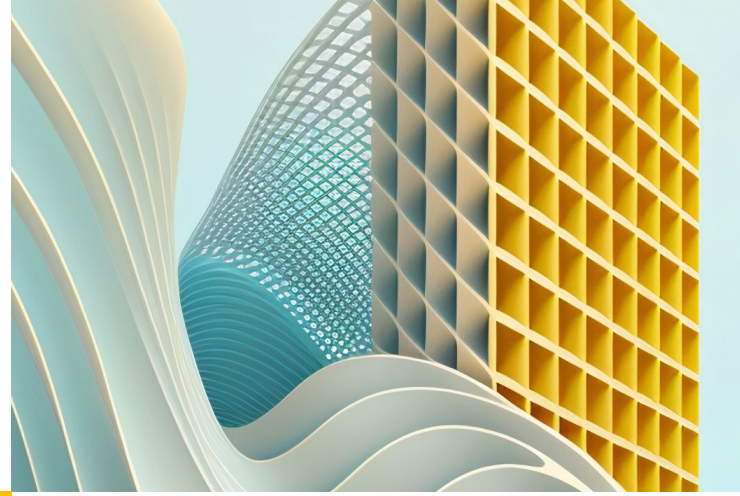
Accelerating Deployment with Serverless: A Practical Guide Using Cloud Run and GitHub Actions

Kacper Prusak | September 2024

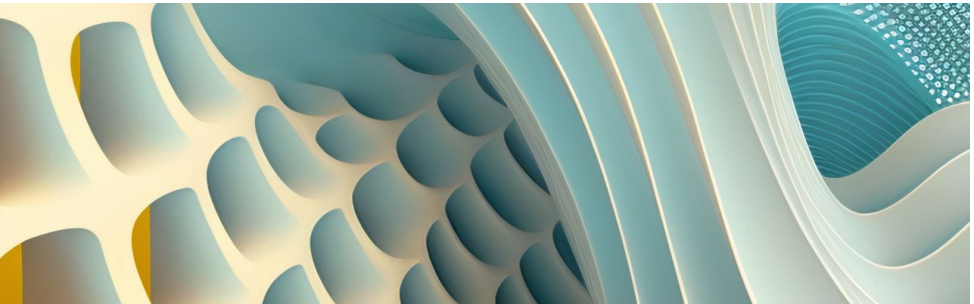
Serverless is a **Cloud Computing Model** where you build and deploy applications without managing the underlying **infrastructure**. The **Cloud Provider** automatically **scales** the application based on **demand**, and you only pay for the **resources used during execution**, eliminating the need for **manual server management**.

The State of The Art.

Deployment processes are now highly **automated** and **Cloud-based**, utilizing **CI/CD**. Applications run in **scalable**, **serverless environments**.



But are we **aware** of it?



Businesses must swiftly adopt to new reality...

... Problem Definition: The challenges of manual deployment, downtime, no scalability, errors, delayed time to market, deployment strategies

Strategy: Automated Deployment using Cloud Run and GitHub Actions.

Cloud Run offers **scalable**, **serverless hosting** for **containerized apps** without managing the **infrastructure**, while **GitHub Actions** streamlines the **CI/CD pipeline**.

Cloud Run - Google Cloud

Cloud Run is a **serverless platform** by **Google Cloud** that allows you to deploy **containerized applications**. It automatically **scales** your app based on **traffic** and abstracts away **infrastructure management**, letting you focus on **code**.


WHAT IF I TOLD YOU...
YOU DON'T HAVE TO CHOOSE
BETWEEN CONTAINERS & SERVERLESS?
imgflip.com

Cloud Run for the Win!

Pros

- **Managed & Serverless**: No need to manage **servers** or worry about **scaling** - **Cloud Run** handles everything.
- **Pay-Per-Use**: You only pay for what you use, leading to **cost savings** for variable traffic.
- **Easy Container Deployment**: Simple to deploy **containerized applications** without worrying about the underlying **infrastructure**.

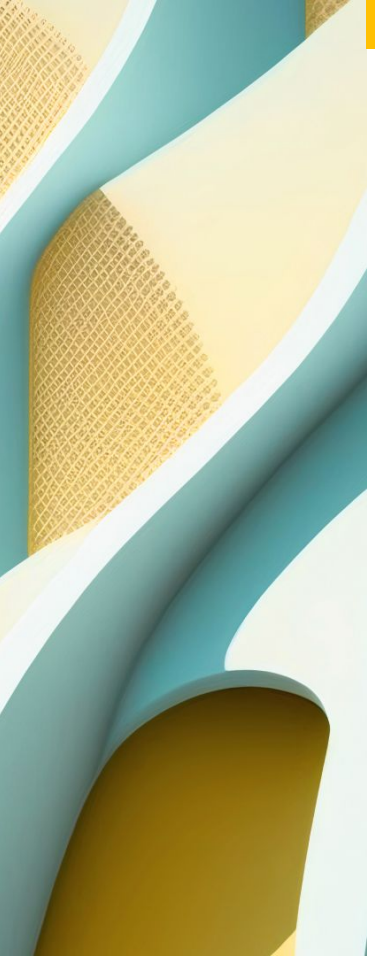
Cons

- **Limited Customization**: Less control over the **OS** and environment compared to **VMs**.
- **Cold Start Latency**: **Cloud Run** can experience delays when scaling from **zero**, affecting **response times**.
- **Not Ideal for Long-Running Tasks**: Designed for short, **stateless** operating; long-running workloads are better suited for **VMs**.

FaaS and Furious by Forrest Brazeal

A CLOUD GURU

The two tribes regarded each other suspiciously in the glow of their brightly blazing production environments.



Cloud Run - Our Roadmap

Service and Revisions

Services refer to the running instances of an application in a cloud environment, typically managing aspects like scaling and traffic routing. Revisions represent specific versions of a service, capturing the configuration and code at a point in time. Each revision allows for easy rollback and version control, enabling developers to track changes and manage deployments effectively.

Secret Management

In each revision of a service, you can reference secrets stored in Secret Manager (GCP) to securely access sensitive information, such as API keys or database credentials. This approach ensures that each deployment can utilize the correct configuration without exposing sensitive data in your code. By managing secrets separately, you enhance security and maintain the flexibility to update sensitive information independently from your application code.

Artifact Management

For storing Docker images, we will use Artifact Registry. This service for secure, scalable storage of container images, providing seamless integration with Cloud Run and facilitating image management and deployment.

Scalability

Cloud Run offers **autoscaling**, so we don't need a **load balancer** in front of our **services**. It automatically scales each **revision** to the required number of **container instances** to handle incoming **requests**. **Cloud Run** can also scale down to zero when there are no **requests**, though this might lead to **cold starts**. To mitigate this, we can set a **minimum number of instances**.

Canary Deployment

A small percentage of traffic is routed to the new version, enabling you to test it in production with real users before a full rollout. If issues arise, you can quickly revert to the previous revision.

Observability

Cloud Monitoring for **Cloud Run** automatically tracks **performance**, **metrics**, and **uptime**, providing **alerts** when thresholds are exceeded. There is no charge for **metrics** on the fully managed version of **Cloud Run**, though **Google Cloud Observability** pricing applies. **Metrics** are automatically captured without setup and can be viewed in both **Cloud Monitoring** and the **Cloud Run console**, with **Cloud Monitoring** offering more advanced **charting** and **filtering options**.

GitHub Actions

GitHub Actions is a **CI/CD platform** natively integrated into GitHub, allowing you to **automate workflows** for building, testing, and deploying code. It responds to GitHub events (e.g., commits, pull requests) and enables easy pipeline creation using **YAML files**. It's designed to streamline **DevOps** and **automation** directly within the GitHub ecosystem.


SETUP-JAVA@V1
SETUP-JAVA@DISCO
SETUP-JAVA@V2
imgflip.com

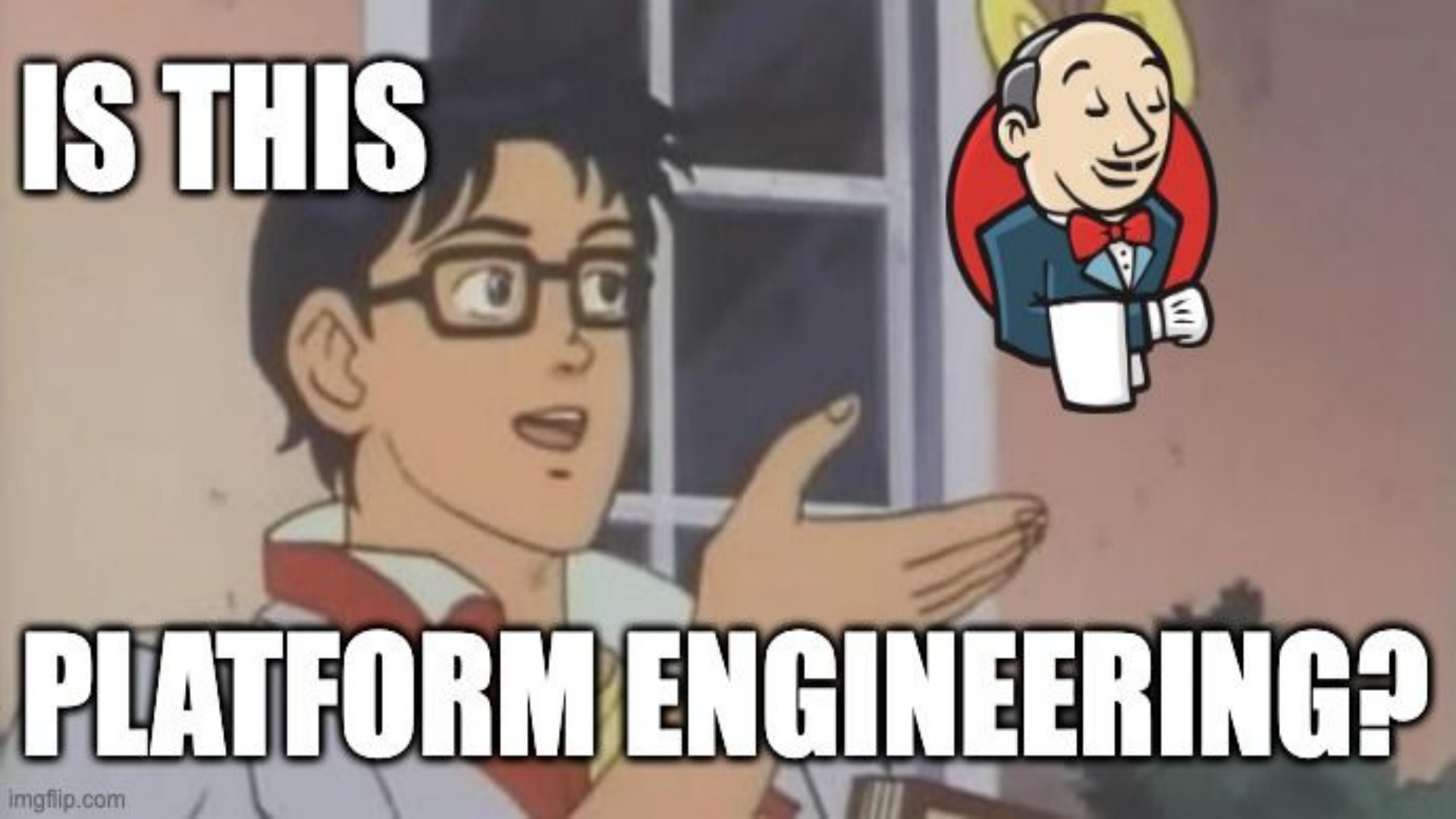
GitHub Actions Against Other CI/CD Solutions

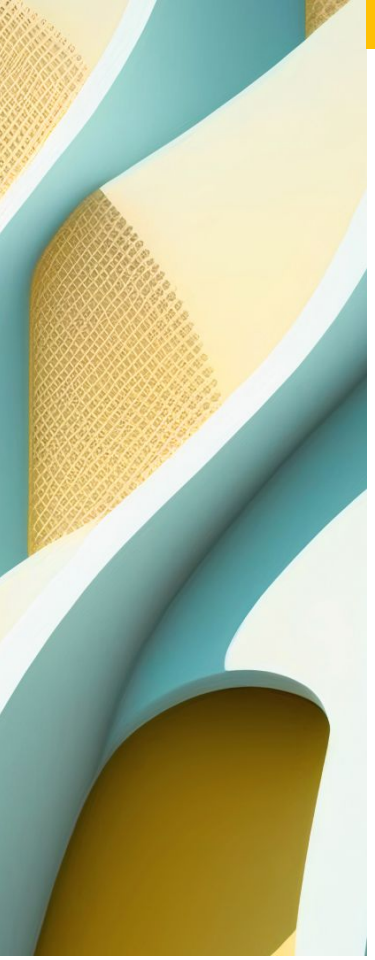
Pros

- **Native GitHub Integration**: Tightly integrated, making it seamless to trigger workflows based on events,
- **Simplified Setup**: Hosted and managed by GitHub, no need to maintain dedicated servers or manage infrastructure (runners),
- **Built-in Marketplace**: Many reusable workflows actions workflows and actions that can be easily integrated into pipeline

Cons

- **Limited Flexibility in Custom Infrastructure:** Actions is hosted, so customizing is more limited.
- **Cost Considerations**: Actions offers free minutes, but might incur extra costs for larger-scale use.
- **Complex Pipeline Management**: It lacks the extensive UI and dashboard management features


IS THIS
PLATFORM ENGINEERING?
imgflip.com



GitHub Actions - Road Map

Secret Storage

Secrets that are used during each **pipeline run** are securely stored in a **GitHub repository**. They are not **hardcoded** into the **code**.

No Runner Setup

In each revision of a service, you can reference secrets stored in Secret Manager (GCP) to securely access sensitive information, such as API keys or database credentials. This approach ensures that each deployment can utilize the correct configuration without exposing sensitive data in your code. By managing secrets separately, you enhance security and maintain the flexibility to update sensitive information independently from your application code.

Verified Tools in Marketplace

GitHub Marketplace provides a selection of verified tools that integrate smoothly into your workflows, ensuring quality and security. GCP tools are maintained by Google and third-party developers, offering regular updates and security patches. This allows you to avoid boilerplate code, streamlining your deployment pipelines and enabling you to focus on building features.

No Integration Configuration

As experts in agile-co-innovation, we embed ourselves in the client teams. We participate early on and contribute to architecture discussions and bring in our institutional knowledge to the table.

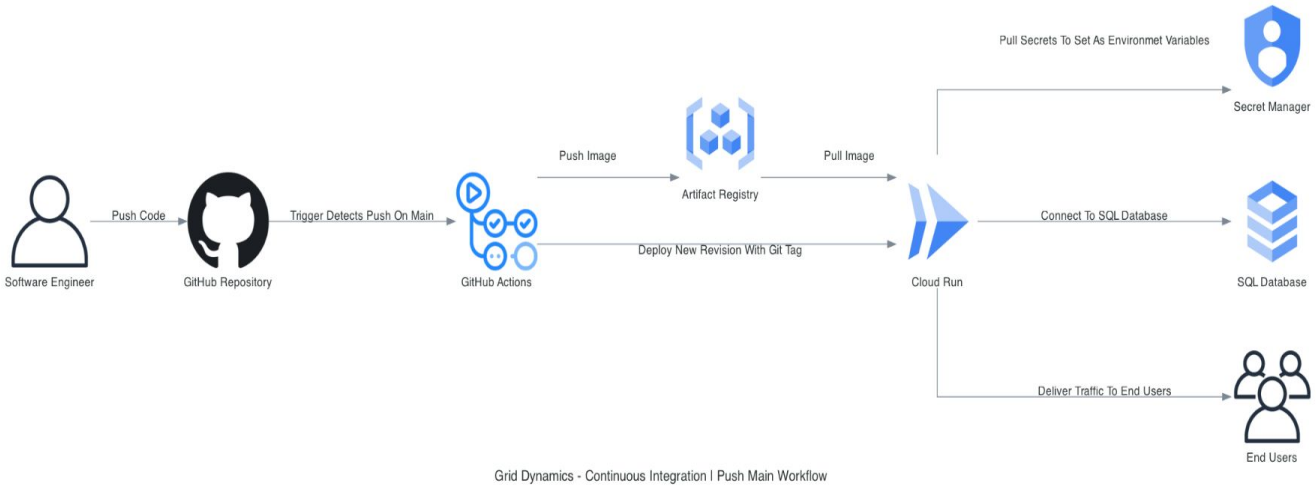
abc

abc

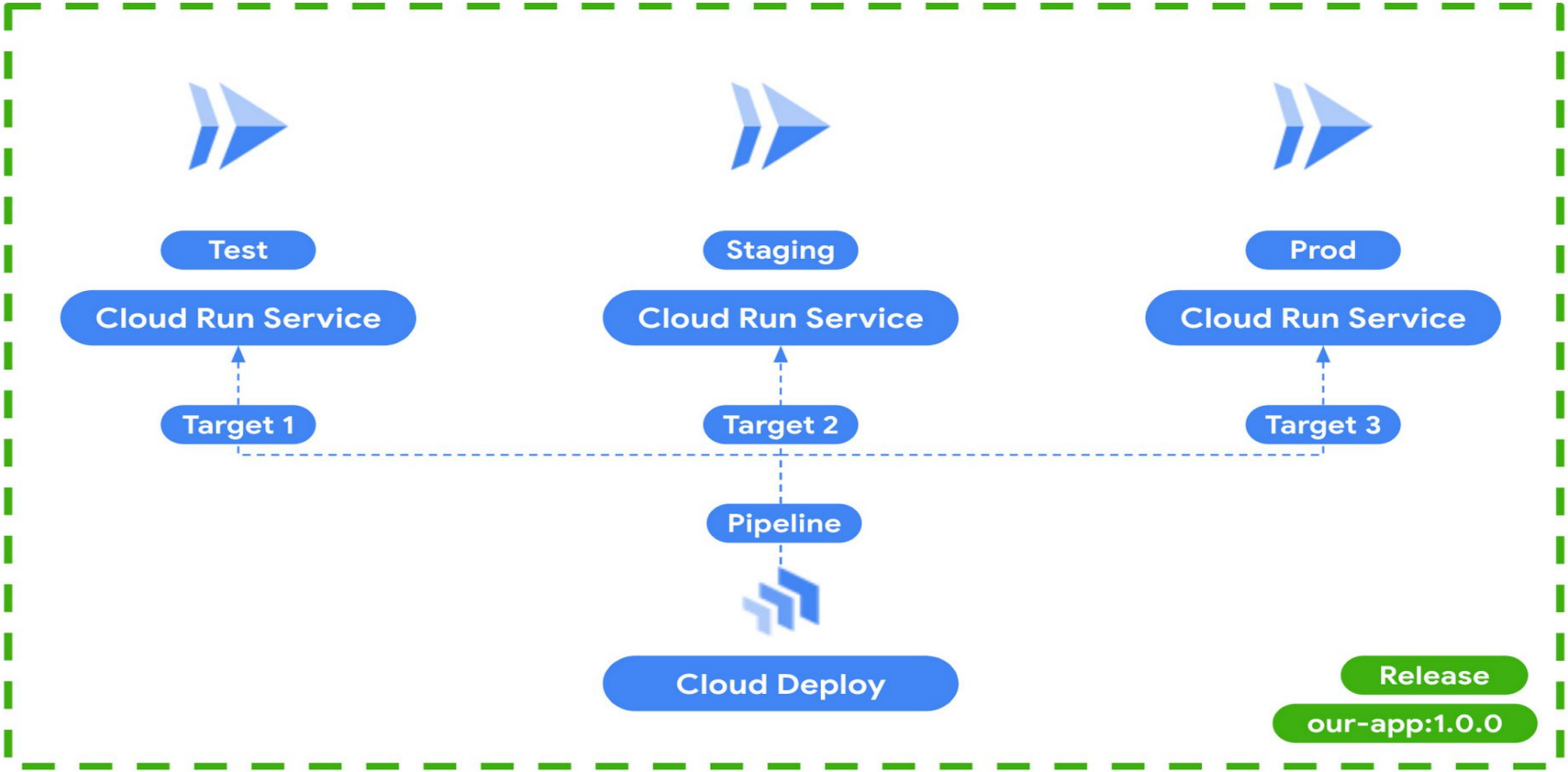
abc

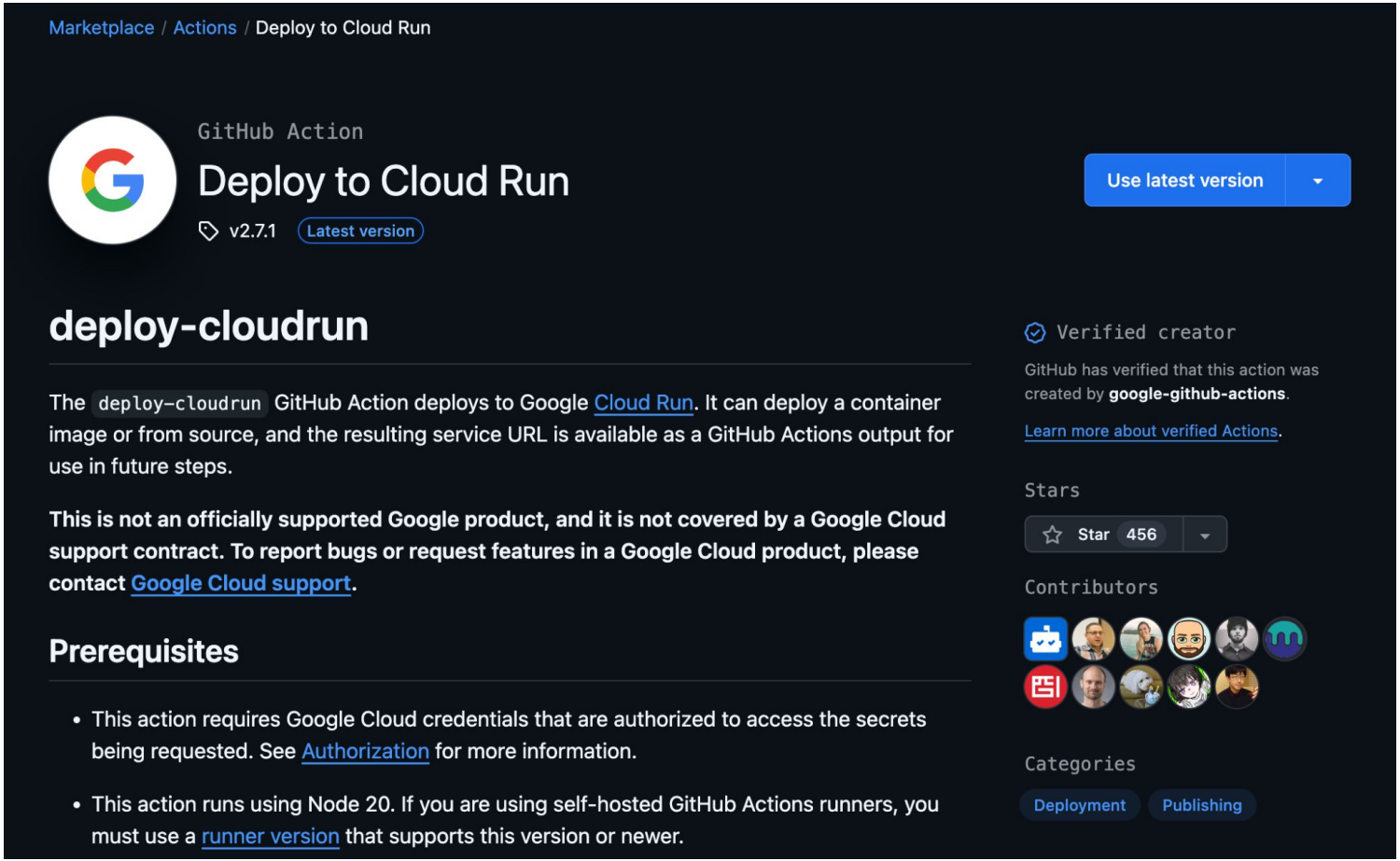
abc

Continuous Integration Repository - Main Push



Grid Dynamics - Continuous Integration I Push Main Workflow


Test
Cloud Run Service
Target 1
Staging
Cloud Run Service
Target 2
Prod
Cloud Run Service
Target 3
Pipeline
Cloud Deploy
Release
our-app:1.0.0


Marketplace / Actions / Deploy to Cloud Run
GitHub Action
Deploy to Cloud Run
v2.7.1
Latest version
Use latest version
deploy-cloudrun
The deploy-cloudrun GitHub Action deploys to Google Cloud Run. It can deploy a container image or from source, and the resulting service URL is available as a GitHub Actions output for use in future steps.
This is not an officially supported Google product, and it is not covered by a Google Cloud support contract. To report bugs or request features in a Google Cloud product, please contact Google Cloud support.
Prerequisites
This action requires Google Cloud credentials that are authorized to access the secrets being requested. See Authorization for more information.
This action runs using Node 20. If you are using self-hosted GitHub Actions runners, you must use a runner version that supports this version or newer.
Verified creator
GitHub has verified that this action was created by google-github-actions.
Learn more about verified Actions.
Stars
Star 456
Contributors
Categories
Deployment
Publishing

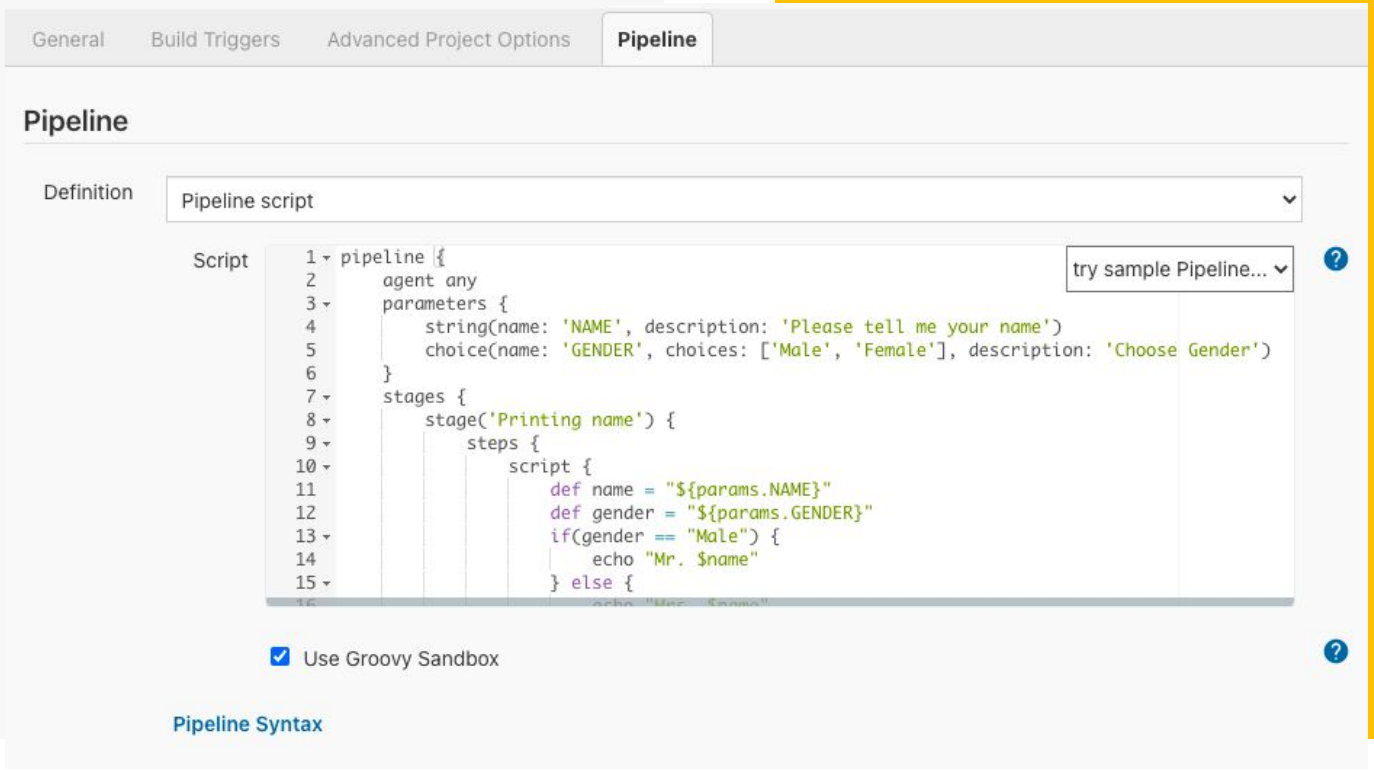
Manual Deployments

Wspomnieć o tym

Tagging Process

Monitoring

Cloud Monitoring for **Cloud Run** automatically tracks **performance**, **metrics**, and **uptime**, providing **alerts** when thresholds are exceeded. There is no charge for **metrics** on the fully managed version of **Cloud Run**, though **Google Cloud Observability** pricing applies. **Metrics** are automatically captured without setup and can be viewed in both **Cloud Monitoring** and the **Cloud Run console**, with **Cloud Monitoring** offering more advanced **charting** and **filtering options**.


General
Build Triggers
Advanced Project Options
Pipeline
Pipeline
Definition
Pipeline script
Script
try sample Pipeline...
1 pipeline {
2 agent any
3 parameters {
4 string(name: 'NAME', description: 'Please tell me your name')
5 choice(name: 'GENDER', choices: ['Male', 'Female'], description: 'Choose Gender')
6 }
7 stages {
8 stage('Printing name') {
9 steps {
10 script {
11 def name = "${params.NAME}"
12 def gender = "${params.GENDER}"
13 if(gender == "Male") {
14 echo "Mr. $name"
15 } else {
Use Groovy Sandbox
Pipeline Syntax

Demo Session

In this part of our meeting I will present a **demo** of my **solution**! :)

Satisfaction

Am I satisfied with my **solution**? Yes! I am satisfied, first of all, because I used an **out-of-the-box solution**, replacing the classic process, shortening it to fewer steps, reducing the **time** and **computing resources** used.