Module 4

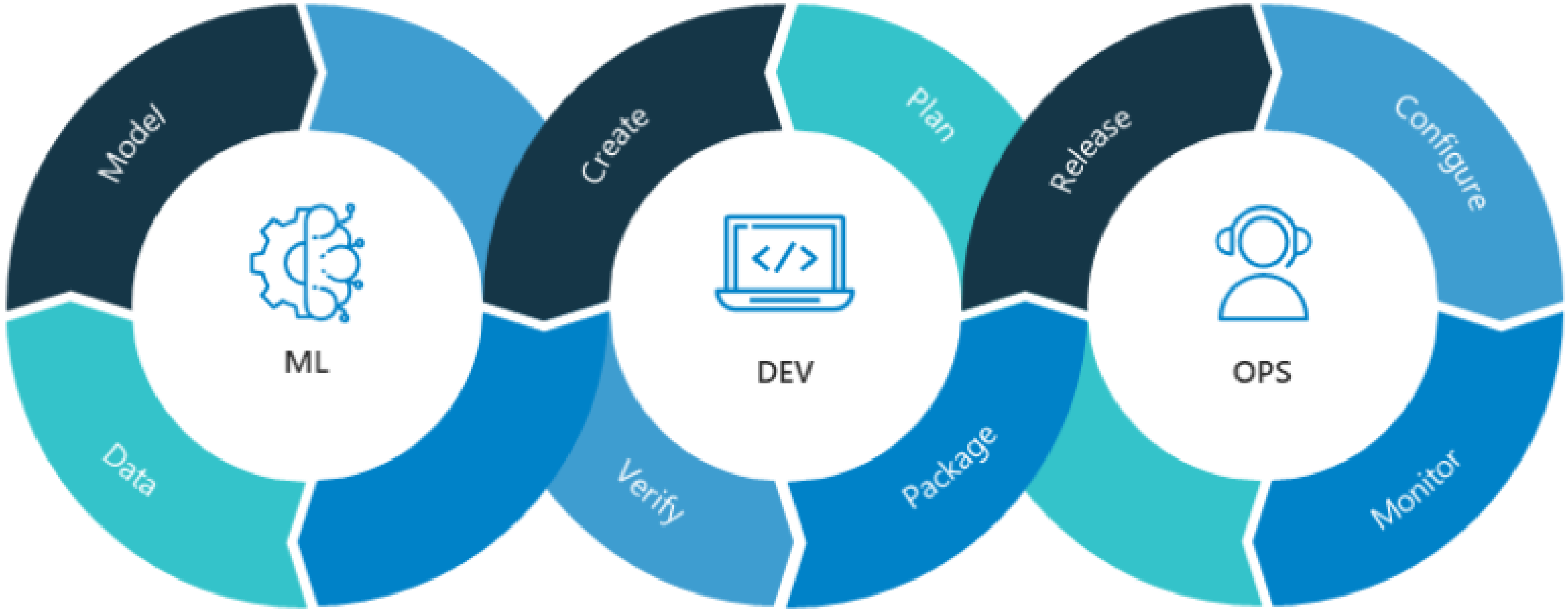
Introduction to CI/CD

CI/CD in MLOps

Psitron

- ✓ In MLOps, CI/CD can be used to automate the process of **building**, **training**, and **deploying machine learning models.**

- ✓ Automating the process of **building**, **training**, and **deploying machine learning models** allows teams to improve the **speed**, **quality**, and **reliability** of their machine learning models.



In MLOps, CI/CD can be used to automate the following steps:

Data preparation:

- ✓ Automate the process of **collecting**, **cleaning** and **pre-processing data** for **training** and **evaluating** the model

Model Development:

- ✓ **Automatically build** and **compile the code changes**, including the model code, as soon as they are committed to the code repository

Model Training:

- ✓ **Automatically train** the **model** using the **prepared data** once the **build process** is done.

Model Evaluation:

- ✓ **Automatically evaluate** the **trained model** using a variety of **evaluation metrics** and **methods**, such as **accuracy, F1-score, and cross-validation.**

Model Validation:

- ✓ **Automatically validate** the results of the **evaluation** to ensure that the **model meets the required standards** and **specifications**

Model Deployment:

- ✓ **Automatically deploy the model** to a **production environment**, such as a **cloud-based environment** or a **hardware-based environment**, once it has passed the **evaluation** and **validation** process.

Model Monitoring:

- ✓ **Automatically monitor** the **deployed model for performance** and **stability** to ensure that the **model is working as expected**.

Model Rollback:

- ✓ **Automatically roll back** to a **previous version** of the model in case of any **issues that may arise.**

CI/CD challenges in Machine Learning

Data versioning

- ✓ ML models heavily depend on data
- ✓ Ensuring the correct version of data is used during model training, testing, and deployment can be challenging
- ✓ A good data versioning strategy is required
- ✓ The ability to reproduce results is necessary

Reproducibility

- ✓ Reproducing model training results by other team members or different environments can be difficult
- ✓ Strict documentation, versioning and testing practices are required

CI/CD challenges in Machine Learning

Model management

- ✓ Managing multiple ML models, their versions, and their dependencies can be difficult
- ✓ A good model management strategy is required
- ✓ A model registry is necessary
- ✓ The ability to roll back to previous versions is necessary

Resource constraints

- ✓ ML models can require large amounts of computational resources and storage
- ✓ Deployment in certain environments can be challenging
- ✓ Careful resource planning and management is required

CI/CD challenges in Machine Learning

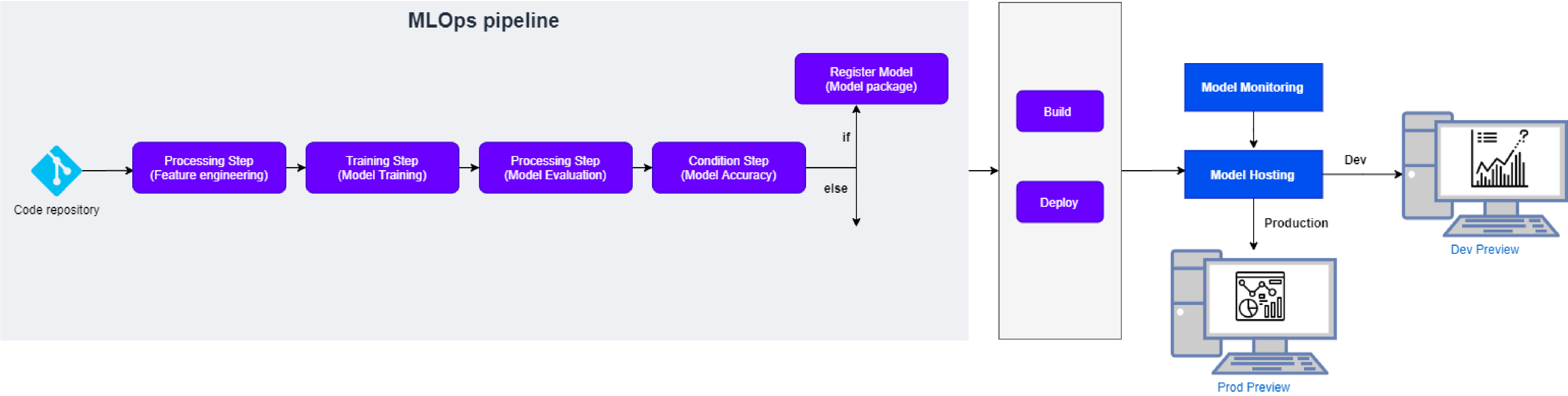
Deployment

- ✓ Deploying ML models in production can be challenging
- ✓ Complexity of the models and ensuring they work correctly in production environment is necessary
- ✓ Good testing and monitoring practices are required

Explainability

- ✓ Machine learning models can be complex and hard to understand
- ✓ This can make it difficult to explain the model's decisions to stakeholders
- ✓ Good model explainability and interpretability practices are required

Steps involved in the CI/CD implementation in ML lifecycle and workflow



MLOps pipeline
Code repository
Processing Step (Feature engineering)
Training Step (Model Training)
Processing Step (Model Evaluation)
Condition Step (Model Accuracy)
if
else
Register Model (Model package)
Build
Deploy
Model Monitoring
Model Hosting
Dev
Production
Dev Preview
Prod Preview

AWS DevOps

1 Code Pipeline

2 Code Commit

3 Code Build

4 Code Deploy

AWS CodeCommit

Securely host highly scalable private Git repositories. Collaborate on code.

- **Secure**, **highly scalable**, **managed source control service** that hosts **private Git repositories**.

  - ✓ Use CodeCommit to store anything from **code to binaries**
  - ✓ Supports the **standard functionality of Git**
  - ✓ Works seamlessly with your existing **Git-based tools**

Benefits

- **Fully managed**

  Eliminates the need to host, maintain, back up, and scale your own source control servers.

- **Secure**

  - ✓ Encrypts your files in transit and at rest.
  - ✓ CodeCommit is integrated with AWS Identity and Access Management (IAM)

- **High availability**
- **Collaborate on code:** Collaborate on code with teammates via pull requests, branching, and merging
- **Faster development lifecycle**
- **Use your existing tools:** Supports all Git commands and works with your existing Git tools

AWS CodeCommit

Hands-on

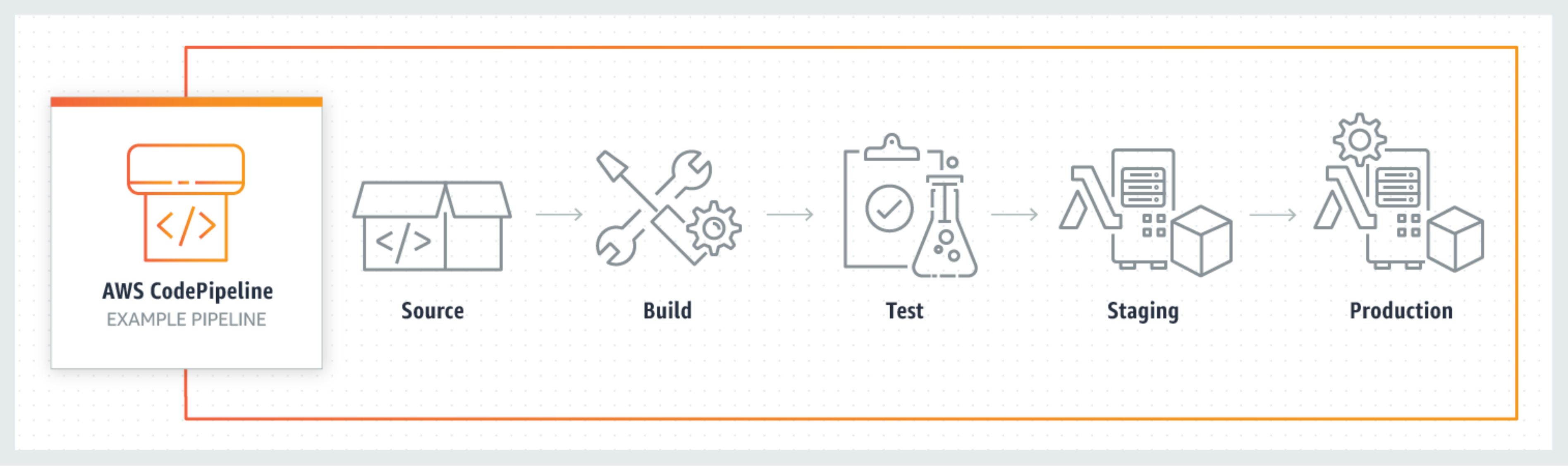
AWS CodePipeline

Automate continuous delivery pipelines for fast and reliable updates

- Fully managed **continuous delivery service**
- Automate your **release pipelines** for **fast and reliable application** and **infrastructure updates**
- Automates the **build**, **test**, and **deploy phases** of your release process every time there is a **code change**, based on the release model you define
- Deliver your features **rapidly** and **reliably**
- Integrate **AWS CodePipeline** with third-party services such as GitHub or with your own custom plugin

AWS CodePipeline

How it works ?



AWS CodeBuild

Build and **test code** with continuous scaling.

- Pay only for the build time you use.
- Fully managed continuous integration service
- **Compiles source code**, **runs tests**, and **produces software packages** that are **ready to deploy**
- No need to **provision**, **manage**, and **scale** your own **build servers**
- Processes multiple builds **concurrently**, so your builds are not left waiting in a queue.
- Charged by the minute for the compute resources you use.

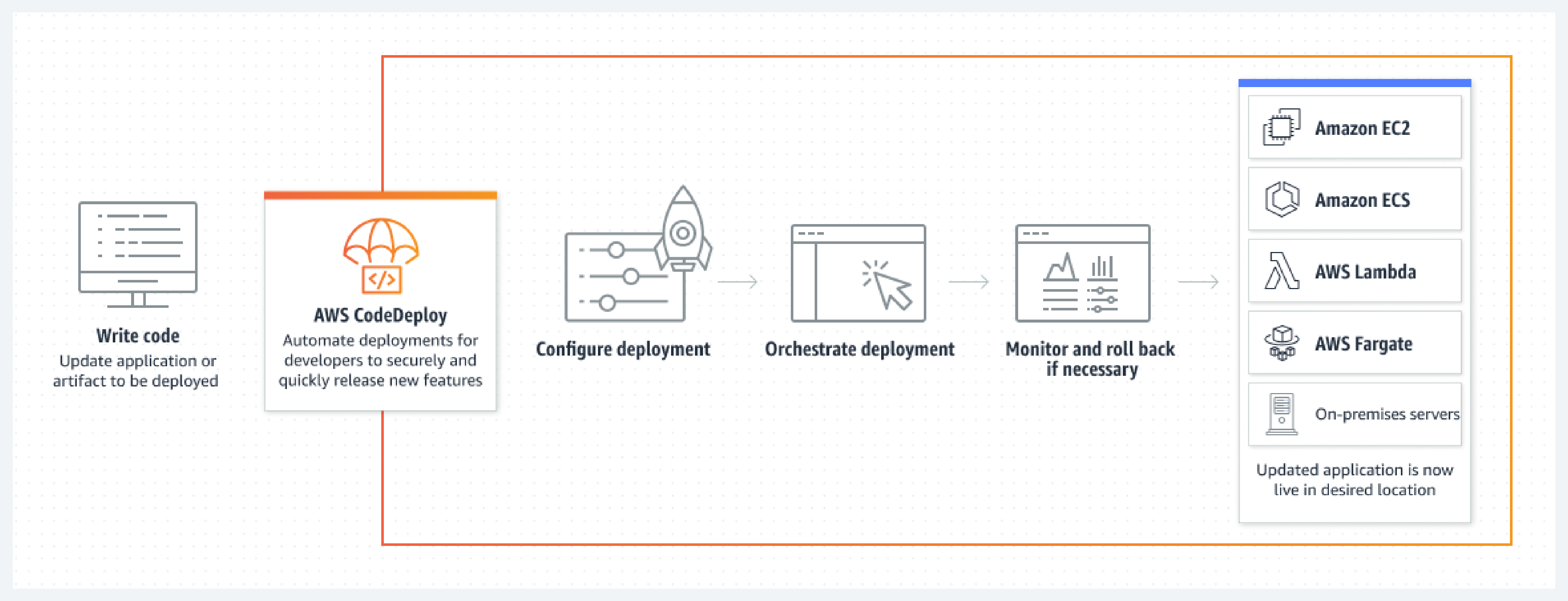
Benefits

- Fully managed build service
- Continuous scaling
- Pay as you go
- Extensible
- Enables continuous integration and delivery
- Secure

AWS CodeDeploy

- ✓ AWS CodeDeploy is a fully managed deployment service
- ✓ Automates software deployments to various compute services, such as **Amazon Elastic Compute Cloud (EC2), Amazon Elastic Container Service (ECS), AWS Lambda** and **on-premises servers**
- ✓ Automates software deployments eliminating manual operations
- ✓ Reduce the error prone manual operations.

AWS CodeDeploy



Write code
Update application or artifact to be deployed
AWS CodeDeploy
Automate deployments for developers to securely and quickly release new features
Configure deployment
Orchestrate deployment
Monitor and roll back if necessary
Amazon EC2
Amazon ECS
AWS Lambda
AWS Fargate
On-premises servers
Updated application is now live in desired location

Benefits:

✓Automated deployments

✓Minimize downtime

✓Centralized control

✓Easy to adopt

AWS CodePipeline

AWS CodeCommit & AWS CodeBuild

Hands-on

AWS CodeBuild

GCP DevOps

Google Cloud

Cloud Source Repositories

✓ **Cloud Source Repositories** is a service provided by GCP for **storing** and **managing source code**

✓ It allows for **collaboration with team members**

✓ Securely stores code with fine-grained access controls

✓ It is integrated with **other GCP tools** such as **Cloud Build**, **Container Registry**, and **Cloud Run** for easy integration with **CI/CD workflow.**

Cloud Source Repositories

Features of Cloud Source Repositories

- ✓ **Multi-language support:** Cloud Source Repositories supports a wide variety of programming languages including **Java**, **Python**, **Go**, and **more**.

- ✓ **Collaboration tools:** Cloud Source Repositories allows for easy collaboration between **team members**, with features such as **code review**, **issue tracking**, and **branch management.**

- ✓ **Secure storage:** Cloud Source Repositories stores code in a **secure**, **encrypted manner**, and provides fine-grained access controls to manage who can view and modify the code.

- ✓ **Integration with other GCP tools:** Cloud Source Repositories can be easily integrated with other GCP tools such as **Cloud Build** and **Container Registry** for **continuous integration** and **deployment**.

Cloud Source Repositories

Cloud Run

- ✓ Cloud Run is a **fully managed compute platform** provided by GCP.
- ✓ It allows developers to **easily deploy** and **run containerized applications.**
- ✓ Cloud Run is built on top of **Kubernetes**
- ✓ It provides **automatic scaling, high availability**, and **logging**.

Cloud Run

key features of Cloud Run

- ✓ Serverless
- ✓ Container-based
- ✓ Automatic scaling
- ✓ High availability
- ✓ Integration with other GCP services

Cloud Run

Cloud Build

- ✓ Code Build is a service provided by Google Cloud Platform (GCP) that enables developers to **build**, **test**, and **deploy their applications** on the GCP

Cloud Build

key features

✓ **Multi-language support:**

Supports building a wide range of languages and runtime environments, including **Java**, **Go**, **Python**, **Node.js**, and more.

✓ **Integrations**

Integrates with other GCP services, such as Container Registry, Kubernetes Engine, and Cloud Storage, which allows for easy deployment of built applications to GCP resources.

Psitron

- ✓ **Configuration file:**
  - Code Build uses a configuration file called `cloudbuild.yaml` to **specify the build steps**
  - Which can include tasks like building a container image, deploying to a Kubernetes cluster, or running tests.

Cloud Build

- ✓ **Pre-defined build steps**

Code Build has pre-made steps that can be used for common tasks like creating container images or deploying apps to App Engine.

- ✓ **Automated testing**

Cloud Build **enables automatic testing** during the **build process**, making it simpler to **identify** and **resolve issues during the early stages** of development.

- ✓ **Continuous integration and continuous delivery (CI/CD)**

Can be integrated with **other CI/CD tools to automate the process of building, testing, and deploying software.**

✓ **Scalability:**

Supports **building** and **testing large codebases**, and can **scale up resources as needed to handle increased build volume.**

✓ **Security:**

Cloud Build uses **build triggers** to ensure that **only authorized users** can start builds and also offers the **option to encrypt sensitive information** in builds.

✓ **Cost-effective**

Charges only for the **compute resources used during builds, making it a cost-effective** option for automating the software development process.

Cloud Build

Cloud Deploy

- ✓ **Google Cloud Deployment Manager** is a GCP service
- ✓ Allows you to **manage cloud resources** using simple configuration file in **YAML or Python**
- ✓ Allows you to **create**, **update** and **delete** resources on GCP.
- ✓ Cloud Deployment Manager can be used to deploy a variety of GCP resources such as **Google Cloud Storage**, **Cloud SQL**, **Cloud Spanner**, **App Engine** and **Kubernetes Engine**.
- ✓ It allows **creating a single configuration file** to describe all the resources to be **deployed** and **managed**.
- ✓ Cloud Deployment Manager allows the use of templates
- ✓ Templates are reusable configurations that can be customized with variables

Cloud Deploy

Artifacts Registry

✓ Google Cloud Artifact Registry is a fully-managed service on GCP.

✓ Allows you to **store**, **manage**, and access **container images** and package types such as **Maven** and **npm packages**.

✓ It is fully integrated with GCP services like Cloud Build, Kubernetes Engine, and Cloud Functions.

✓ It makes it easy to build, deploy and manage applications on GCP.

Artifacts Registry

key features of Cloud Artifact Registry

- Multi-format support
- Scalable and highly available
- Integrated with GCP services
- Automated vulnerability scanning
- Cost-effective:

App engine

Google App Engine is a fully-managed platform for building and deploying **web applications** and services on the Google Cloud Platform (GCP)

Key features

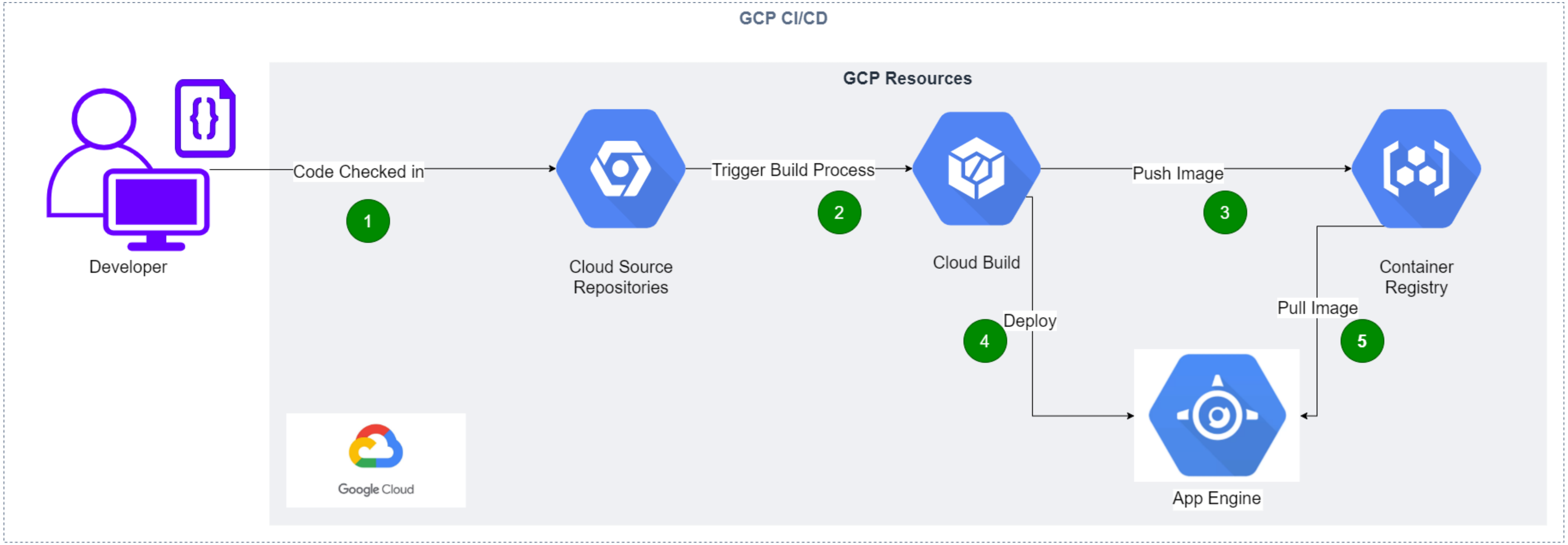
Popular programming languages

Build your application in Node.js, Java, Ruby, C#, Go, Python, or PHP.

Fully managed

A fully managed environment lets you focus on code while App Engine manages infrastructure concerns.

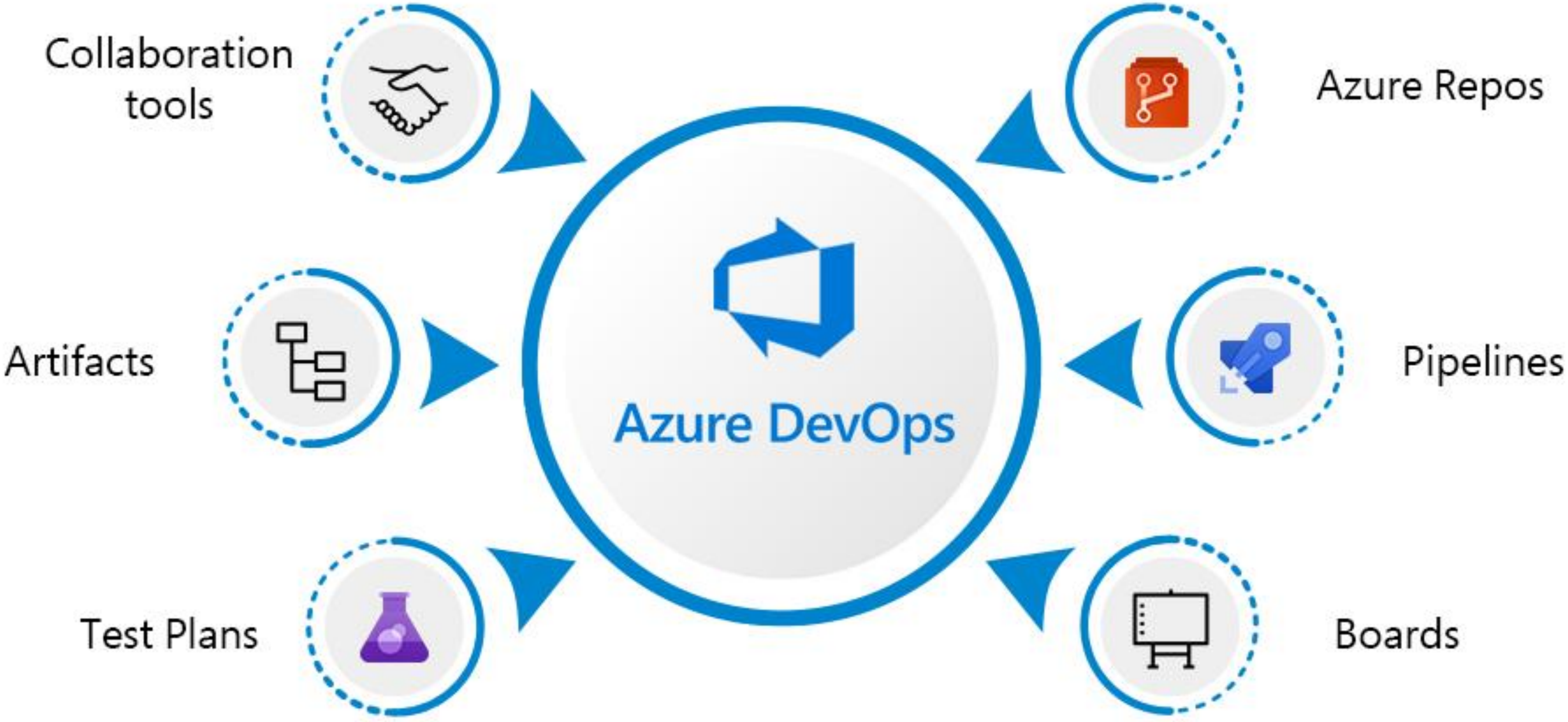
Psitron


GCP CI/CD
GCP Resources
Developer
Code Checked in
1
Cloud Source Repositories
Trigger Build Process
2
Cloud Build
Push Image
3
Container Registry
Deploy
4
Pull Image
5
App Engine
Google Cloud

Hands-on

GCP CI/CD

Azure DevOps



Introduction Azure DevOps

Azure Boards

Deliver value to your users faster using proven agile tools to plan, track, and discuss work across your teams.

Azure Pipelines

Build, test, and deploy with CI/CD that works with any language, platform, and cloud. Connect to GitHub or any other Git provider and deploy continuously.

Azure Test Plans

Test and ship with confidence using manual and exploratory testing tools

Azure Repos

Get unlimited, cloud-hosted private Git repos and collaborate to build better code with pull requests and advanced file management

Azure Artifacts

Create, host, and share packages with your team, and add artifacts to your CI/CD pipelines with a single click.

Azure Pipelines

Cloud-hosted pipelines for Linux, Windows and macOS, with unlimited minutes for open source

Any language, any platform, any cloud

✓ Build test, and deploy Nodejs, Python, Java, PHP Ruby, C/C++, .NET, Android, and iOS apps. Run in parallel on Linux, macOS, and Windows. Deploy to Azure, AWS GCP or on-premises

Extensible

✓ Explore and implement a wide range of community-built, build, test, and deployment tasks, along with hundreds of extensions from Slack to SonarCloud Support for YAML reporting and more" simplify this sentnces

Containers and Kubernetes

✓ Easily build and push images to container registries like Docker Hub and Azure Container Registry. Deploy containers to individual hosts or Kubernetes.

Azure Pipelines

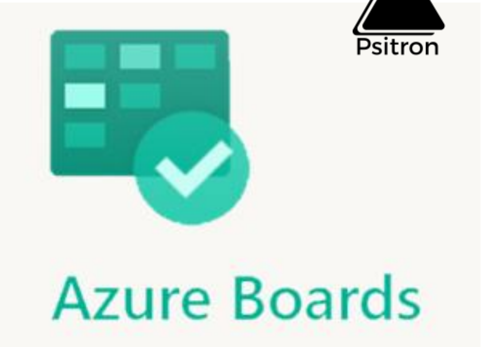
Cloud-hosted pipelines for Linux, Windows and macOS, with unlimited minutes for open source

Best-in-class for open source

Ensure fast continuous integration/continuous delivery (CVCD) pipelines for every open source project. Get unlimited build minutes for all open source projects with up to 10 free parallel jobs across Linux, macOS and Windows

Azure Boards

Track work with Kanban boards, backlogs, team dashboards & custom reporting



Connected from idea to release:

Track all your ideas at every development stage and keep your team aligned with all code changes linked directly to work items.

Scrum ready

Use built-in scrum boards and planning tools to help your teams run sprints, stand-ups & planning meetings.

Project insights

Gain new insights into the health and status of your projects with powerful analytics tools and dashboard widgets

Azure Repos

Psitron

Unlimited private Git repo hosting and support for TFVC that scales from a hobby project to the world's largest Git repositories

Azure Repos

Works with your Git client

Securely connect with and push code into your Git repos from any IDE, editor or Git client

Web hooks and API integration

Add validations and extensions from that marketplace or build your own using web hook and REST APIs

Semantic code search

Quickly find what you're looking for with code-aware search that understands classes and variables

Azure Test Plans

Get end-to-end traceability. Run tests and log defects from your browser. Track and assess quality throughout your testing lifecycle

Azure Test Plans

Capture rich data

Capture rich scenario data as you execute tests to make discovered defects actionable. Explore user stories without test cases or test steps. You can create test cases directly from your exploratory test sessions.

Test across web and desktop

Test your application where it lives. Complete scripted tests across desktop or web scenarios. Test on-premises application from the cloud and vice-versa.

Get end-to end traceability

Leverage the same test tools across your engineers and user acceptance testing stakeholders. Pay for the tools only when you need them.

Azure Artifacts

Create and share Maven , npm, and NuGet package feeds from public and private sources – fully integrated into CI/CD pipelines

Azure Artifacts

Manage all package types

Get universal artifact management for Maven, npm, and NuGet

Add packages to any pipeline

Share packages, and use build-in CI/CD, versioning and testing.

Share code efficiently

Easily share code across small teams and large enterprices

Psitron

Hands-on

Azure DevOps