

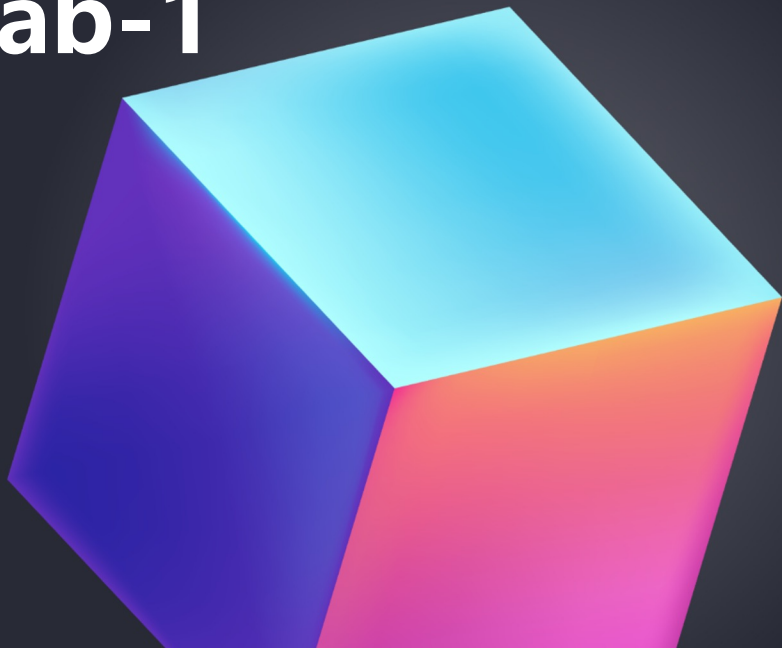


# CCF 中国数据库暑期学校 2023

## PingCAP 工程实践 Lab-1

在 AWS EKS 上部署并操作 TiDB 集群

By 徐一帆



## Instructor Bio

- 徐一帆 (SabaPing@GitHub) 是 PingCAP 自治服务 (Autonomous Service) 团队的后端工程师，该团队负责帮助内部用户和客户舒适地使用 TiDB 和 TiDB Cloud。
- 他涉猎广泛，曾参与 observability-for-TiDB, TiDB-on-Kubernetes, cloud-infrastructure 和 data-streaming 等多个领域的工作。他拥有 Golang 和 Scala 经验。
- 他目前致力于 TiDB Cloud 控制平台的改进。



# Lab Introduction

本实验在 AWS EKS 上部署了一个 TiDB 集群。TiDB 集群由 TiDB-Operator 管理，部署过程使用了 Pulumi 自动化。

- AWS EKS 是一个托管的 Kubernetes 服务  
Kubernetes提供了容器化应用的自动部署、扩展和管理。
- TiDB 是一个开源的分布式 SQL 数据库，支持 HTAP 工作负载。
- 当在 AWS EKS 上部署TiDB集群时，用户可以获得 TiDB 提供的所有功能, 同时利用在托管 Kubernetes 服务上运维 TiDB 集群的好处。



# Learning Objectives

01

## Kubernetes

了解 Kubernetes 和  
AWS EKS 的基本用法

02

## TiDB

了解 TiDB 集群的核心  
特性

03

## TiDB Operator

了解 operator 模式的基本原  
理, 并学习使用 TiDB  
Operator 在 Kubernetes 上  
部署 TiDB 集群

04

## Pulumi

使用 Pulumi IaC 工具  
自动化部署流程

## Tasks and Credits

1. (25 points) 创建 EKS 集群
2. (25 points) 使用 TiDB Operator 部署集群
3. 中场休息
4. (20 points) 探索 TiDB 的使用方式
5. (20 points) 使用 TiDB Operator 为集群扩容
6. (10 points) 清理：销毁 EKS 集群

(20 bonus points) 加分项：基于上面学到知识，调整 TiDB slow-log 阈值，并设计能够触发阈值的 SQL statements

# Introduction

## Introduction

**Cloud computing:** the delivery of computing services over the internet ("the cloud")

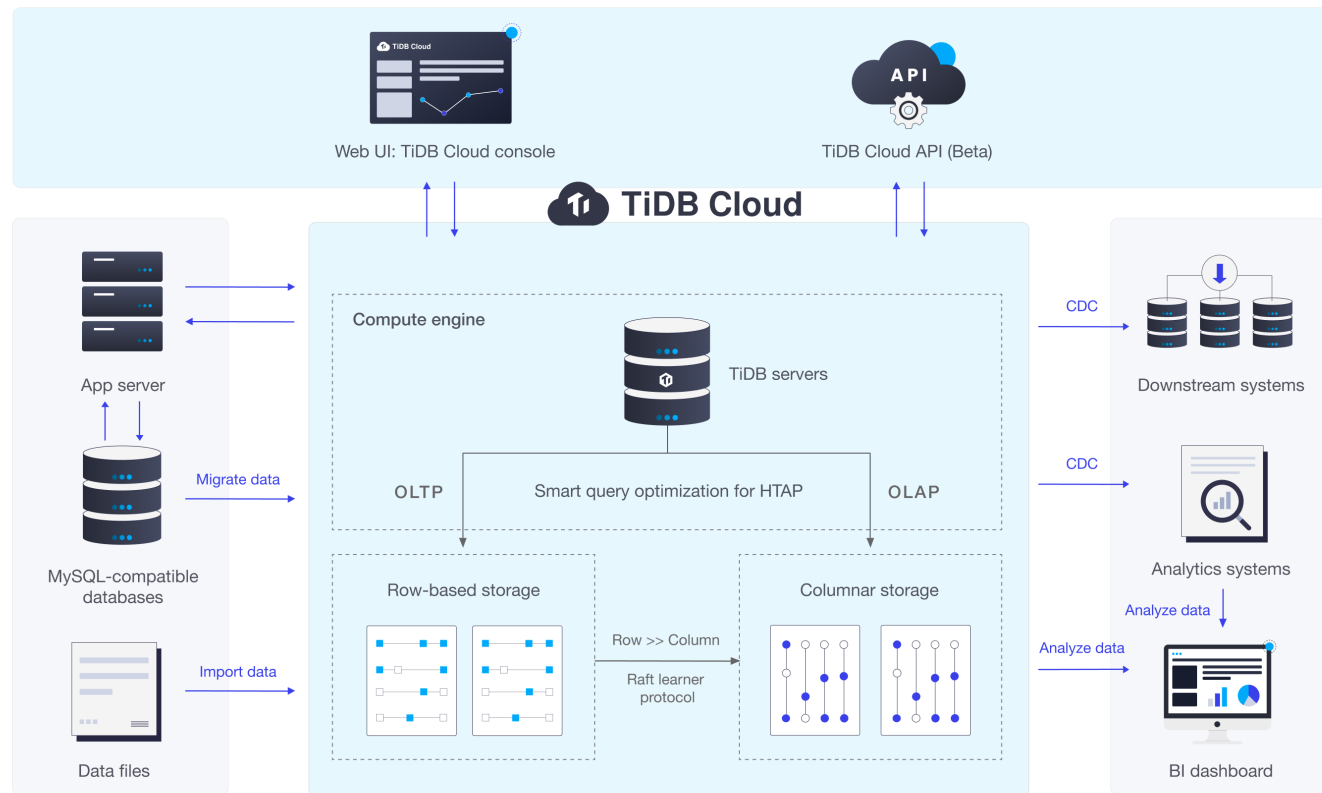
**AWS EKS:** a managed Kubernetes service

**TiDB:** an open-source distributed SQL database that supports HTAP workloads

# Introduction

## TiDB Cloud

At PingCAP,  
we build TiDB Cloud  
based on technologies  
covered in this Lab 1!





# Step 1: Create an EKS Cluster

## Step 1: Create an EKS Cluster

**About AWS EKS:** a robust managed Kubernetes service

**What is Kubernetes:** an open-source container  
orchestration system

**Why container:** Isolation, Portability, Resource efficiency

## Step 1: Create an EKS Cluster

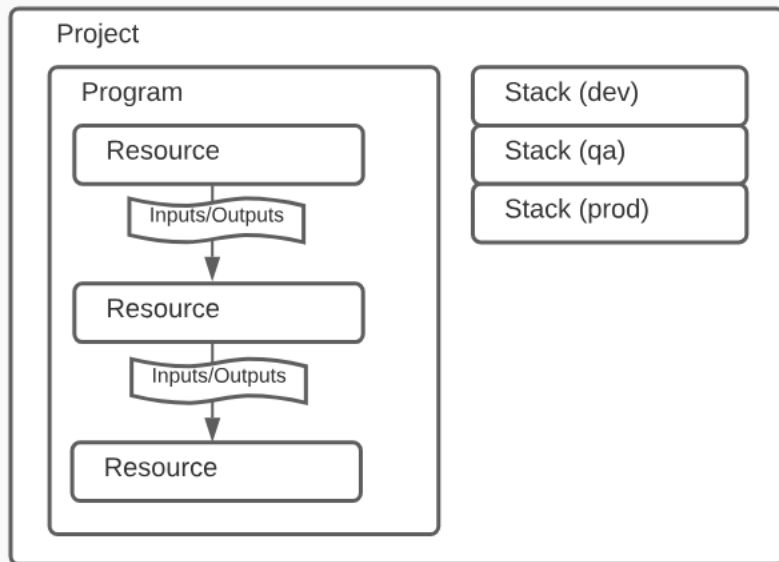
### About Pulumi

**Why Pulumi?** version control, testing, DevOps, high-level

**The Age before IaC:** Darkness!

## Step 1: Create an EKS Cluster

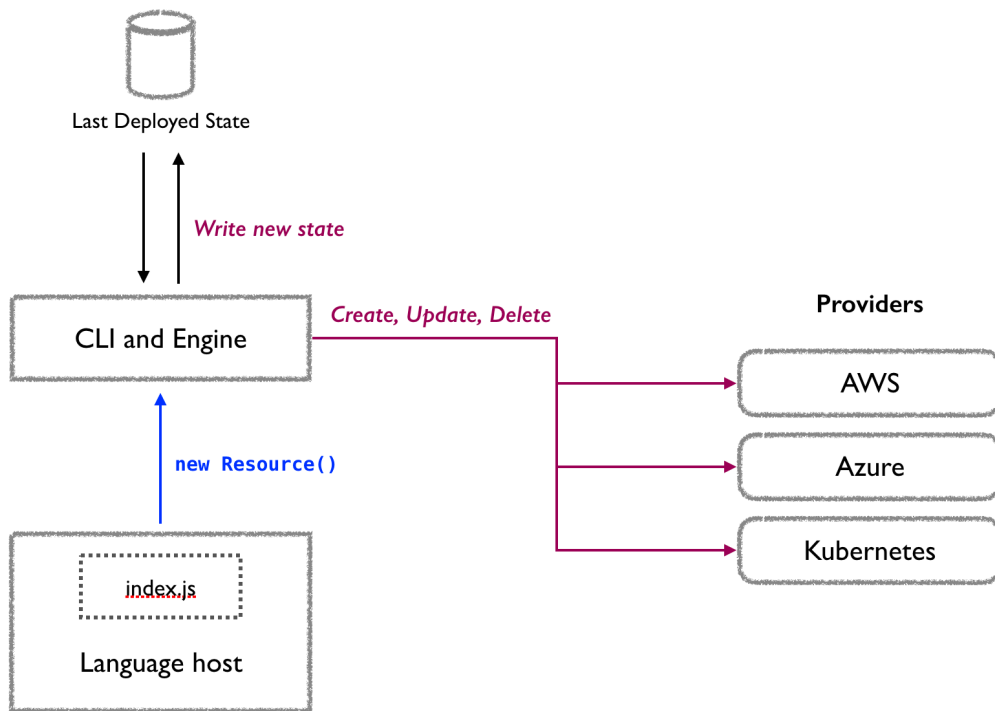
# Pulumi Basic Concepts



## Step 1: Create an EKS Cluster

Pulumi up

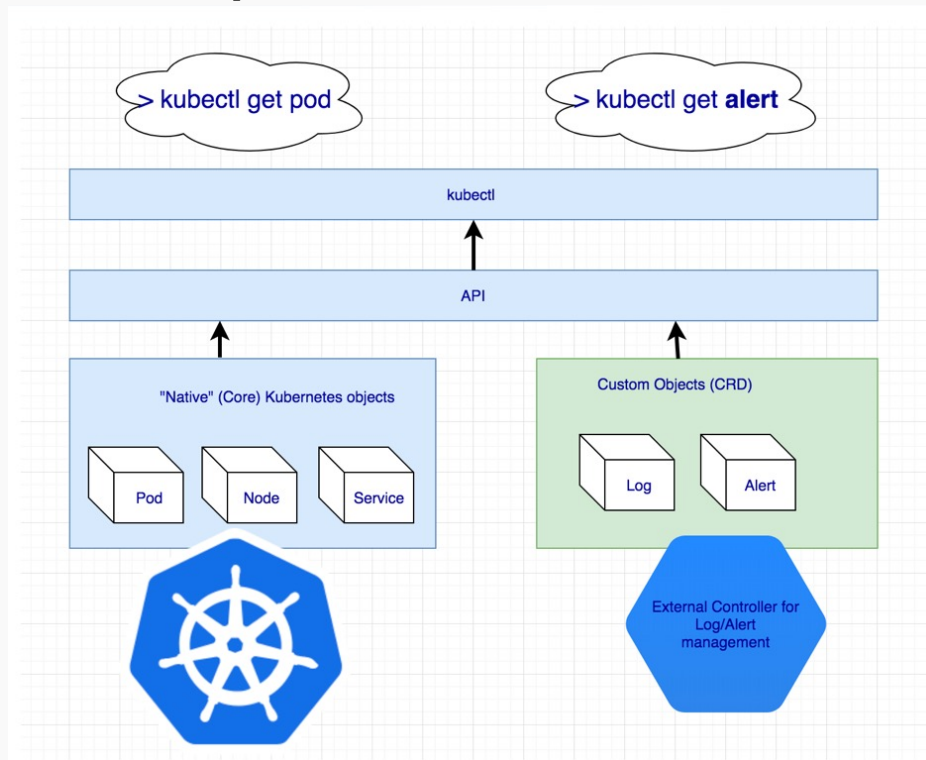
What happened?



## Step 2: Deploy a TiDB Cluster with TiDB Operator

## Step 2: Deploy a TiDB Cluster with TiDB Operator

### What is CRD?



## Step 2: Deploy a TiDB Cluster with TiDB Operator

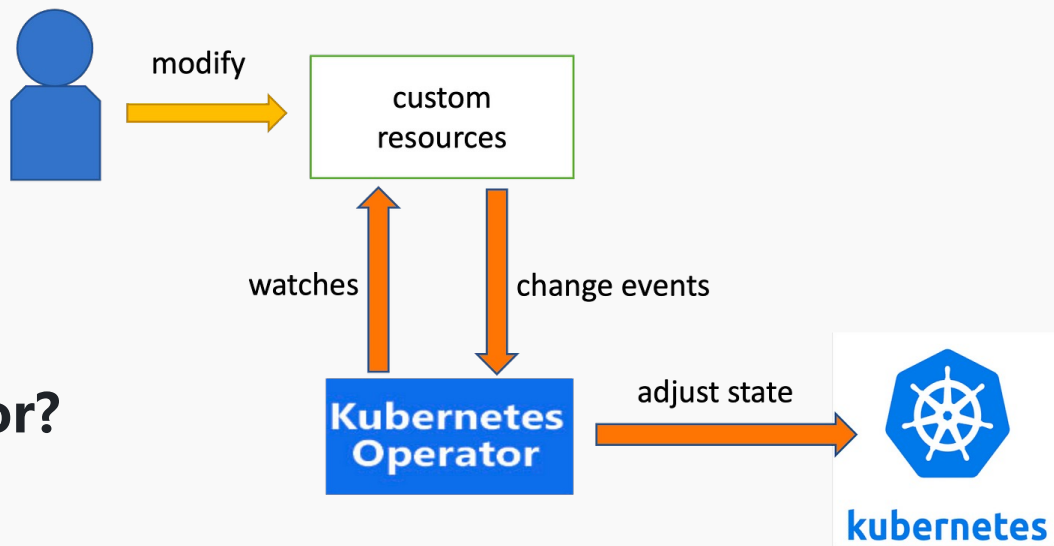
### What is Operator

Software extensions to Kubernetes, the manager of CRD



## Step 2: Deploy a TiDB Cluster with TiDB Operator

How Do We  
Interact with Operator?

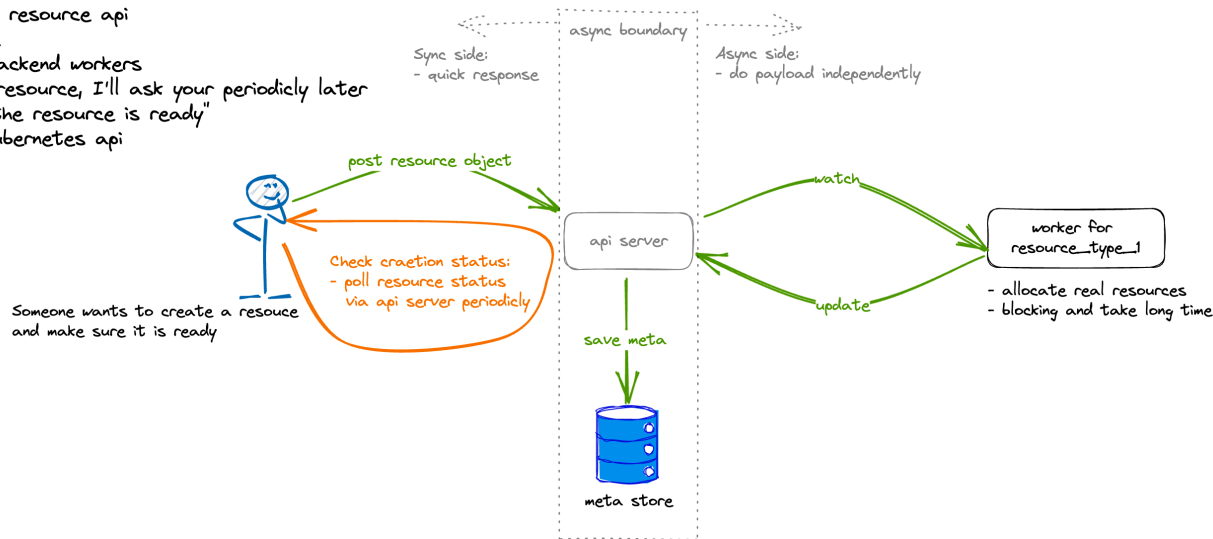


## Step 2: Deploy a TiDB Cluster with TiDB Operator

### How Does Operator Work --- Reconcile Pattern

Reconcile pattern:

- A declarative resource api
- A meta store
- Some async backend workers
- "I want this resource, I'll ask you periodically later for whether the resource is ready"
- Inspired by kubernetes api



## Step 2: Deploy a TiDB Cluster with TiDB Operator

**Do it**

Pulumi up

## Step 2: Deploy a TiDB Cluster with TiDB Operator

### **TiUP(Bare-Metal) vs TiDB Operator(On-Kubernetes)**

TiUP offers simplicity, while TiDB Operator provides cloud-native portability, automation, and ecosystem leverage on Kubernetes.

Break (15 minutes)

## Step 3: Explore TiDB Basic Usage

## Step 3: Explore TiDB Basic Usage

### Why TiDB?

- Scalability
- High-performance analytics with strong consistency and less technical complexity

类型	传统数据库中间件 / 分库分表	TiDB
强一致的分布式事务	不支持	支持
水平扩展	不支持	支持
复杂查询 (JOIN/ GROUP BY/...)	不支持	支持
无人工介入的高可用	不支持	支持
业务兼容性	低	高
多维度支持	不友好	友好
全局 ID 支持	不友好	友好
机器容量	很浪费	按需扩容

## Step 3: Explore TiDB Basic Usage

**Do it**



## Step 4: Scale up TiDB Cluster with TiDB Operator

## Step 4: Scale up TiDB Cluster with TiDB Operator

**Do it**

Pulumi up

## Bonus Point: Config the TiDB Slow-Log Threshold

## Bonus Point: Config the TiDB Slow-Log Threshold

### What is TiDB Slow-log?

TiDB Slow-log is a feature of TiDB that records SQL statements that take longer than a certain threshold to execute.

## **Bonus Point: Config the TiDB Slow-Log Threshold**

### **Why We Need to Adjust Slow-Log Threshold?**

Real-world requirements matter.

## Bonus Point: Config the TiDB Slow-Log Threshold

### What Kinds of SQL Statements May Slow Down?

- Large table scans
- Expensive joins
- Large sorts

## Bonus Point: Config the TiDB Slow-Log Threshold

### Do it

You can refer to the `cheat_scripts.sh` in step folder



**THANK YOU.**

