

2018 THINK IN CLOUD BEIJING

基于云计算构建机器学习系统的实践

UCloud 叶理灯

Directory

目录

01

云计算的发展

02

Serverless 产品及架构

03

机器学习在线推测

04

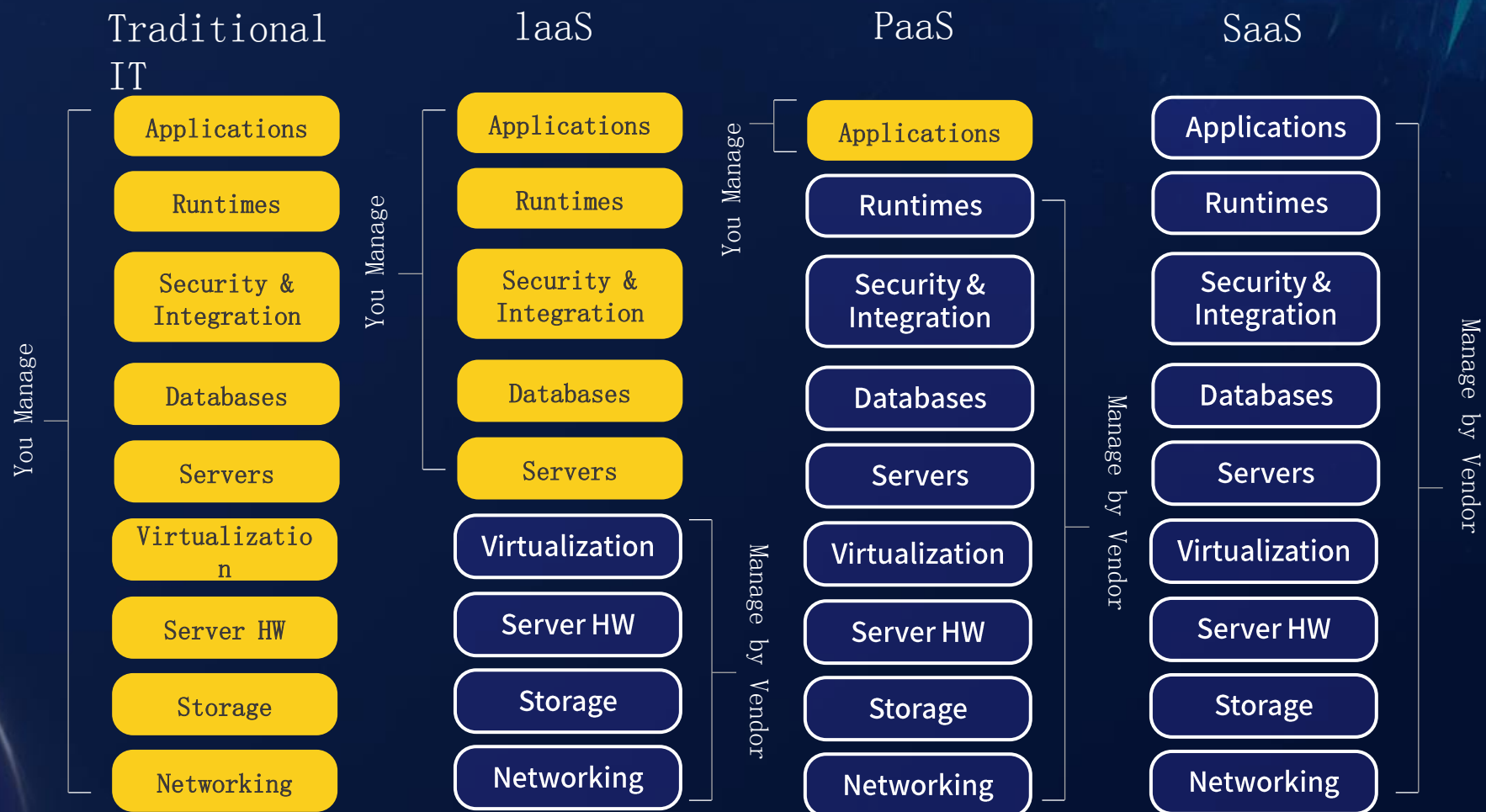
构建在线推测系统

05

应用案例

01

云计算的发展



硬件
托管

运行环
境托管

架构
托管

代码
托管

FaaS



托管
粒度

研发
友好

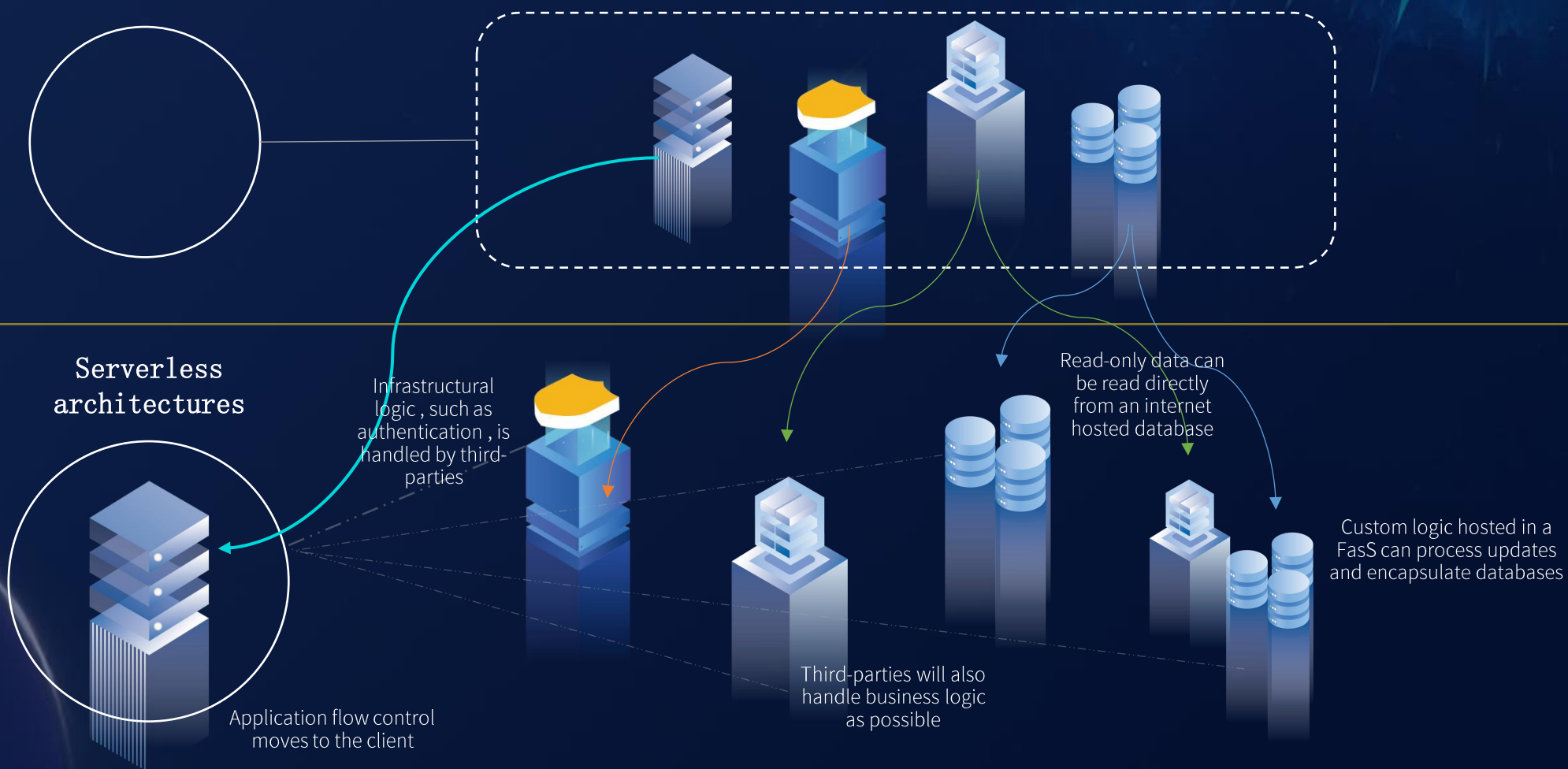
成本

效率

02

Serverless 产品及架构

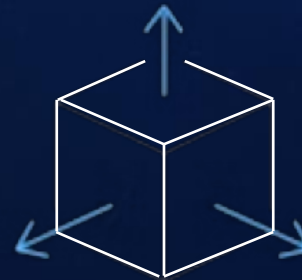
A traditional internet app has a client communicating with a long-lived server process that handles most aspects of the application's logic



Serverless Means...



No Server Management



Flexible Scaling



High Availability

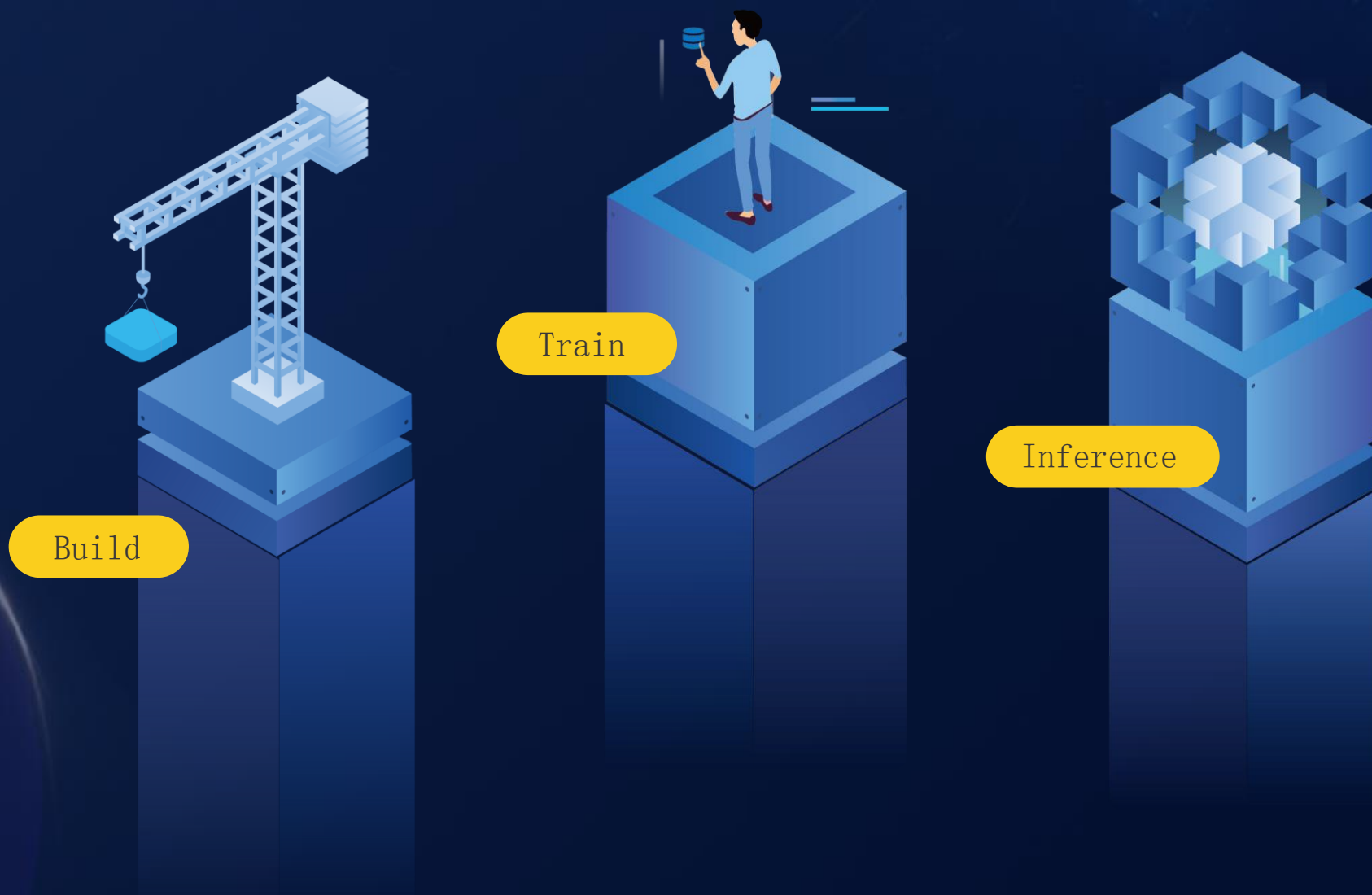


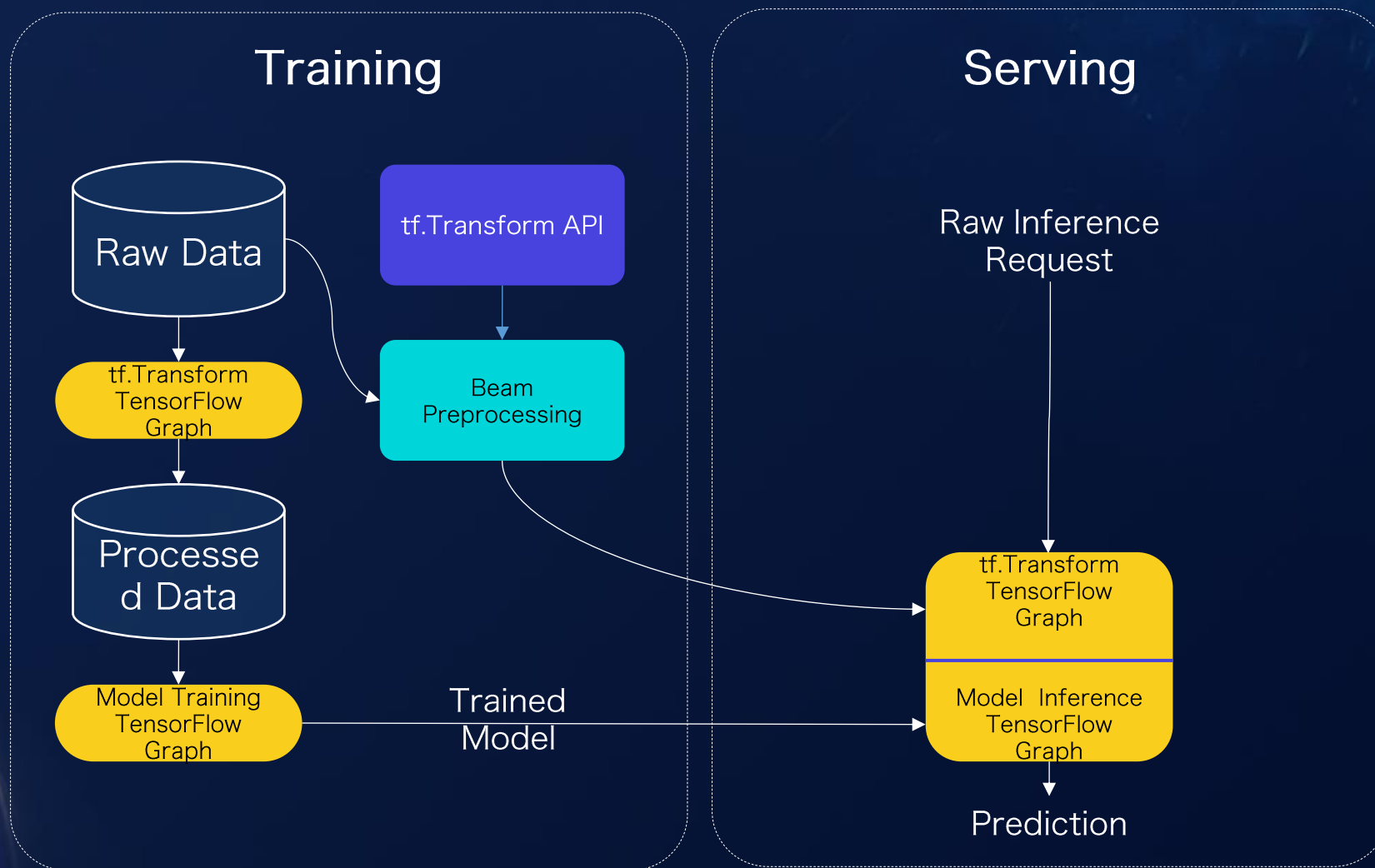
No Idle Capacity

03

机器学习在线推测

机器学习系统





机器学习系统

无状态

Auto
Scaling

高可用

高并发

海量
计算

04

构建在线推测系统

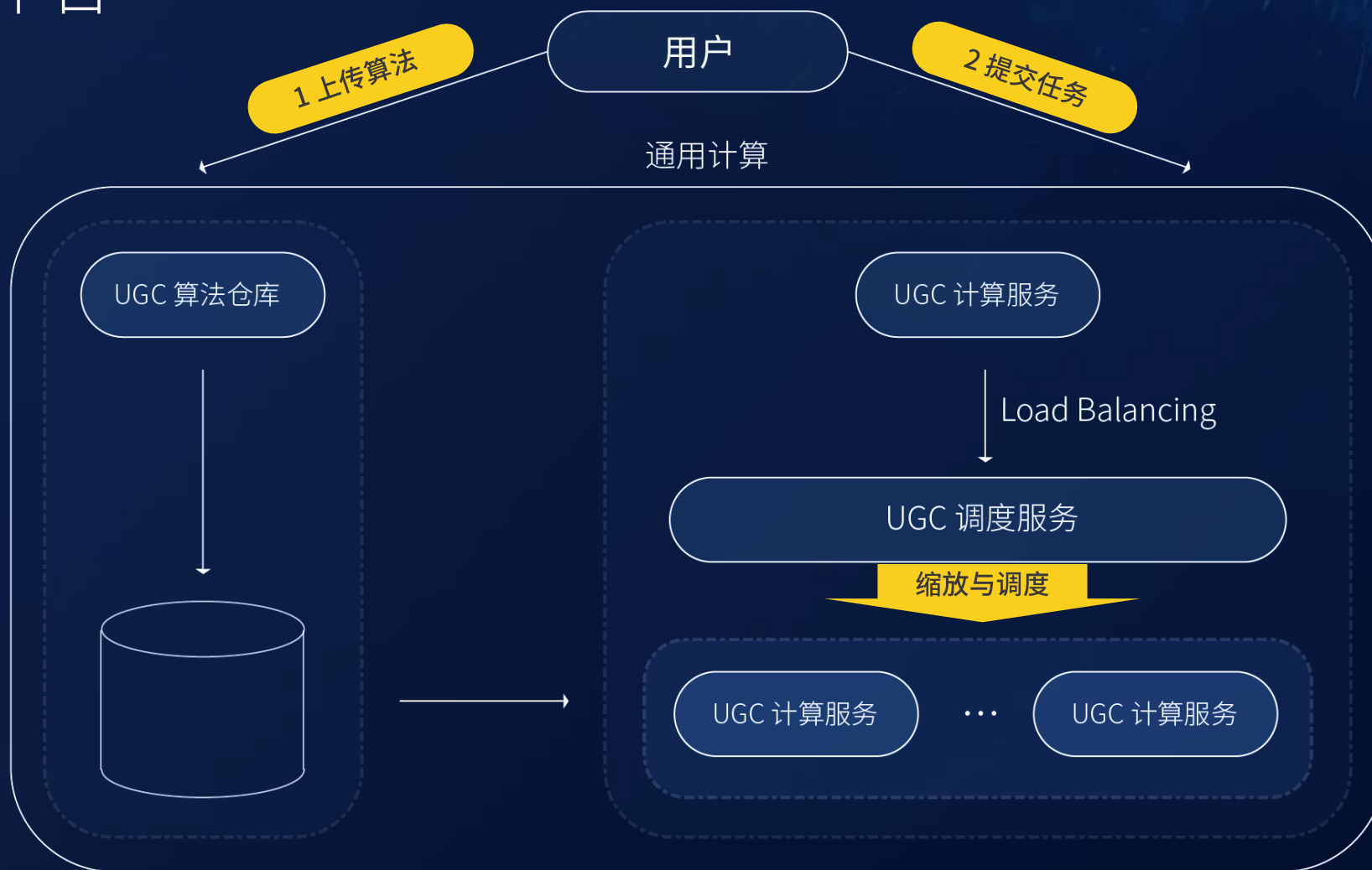


III 计算平台

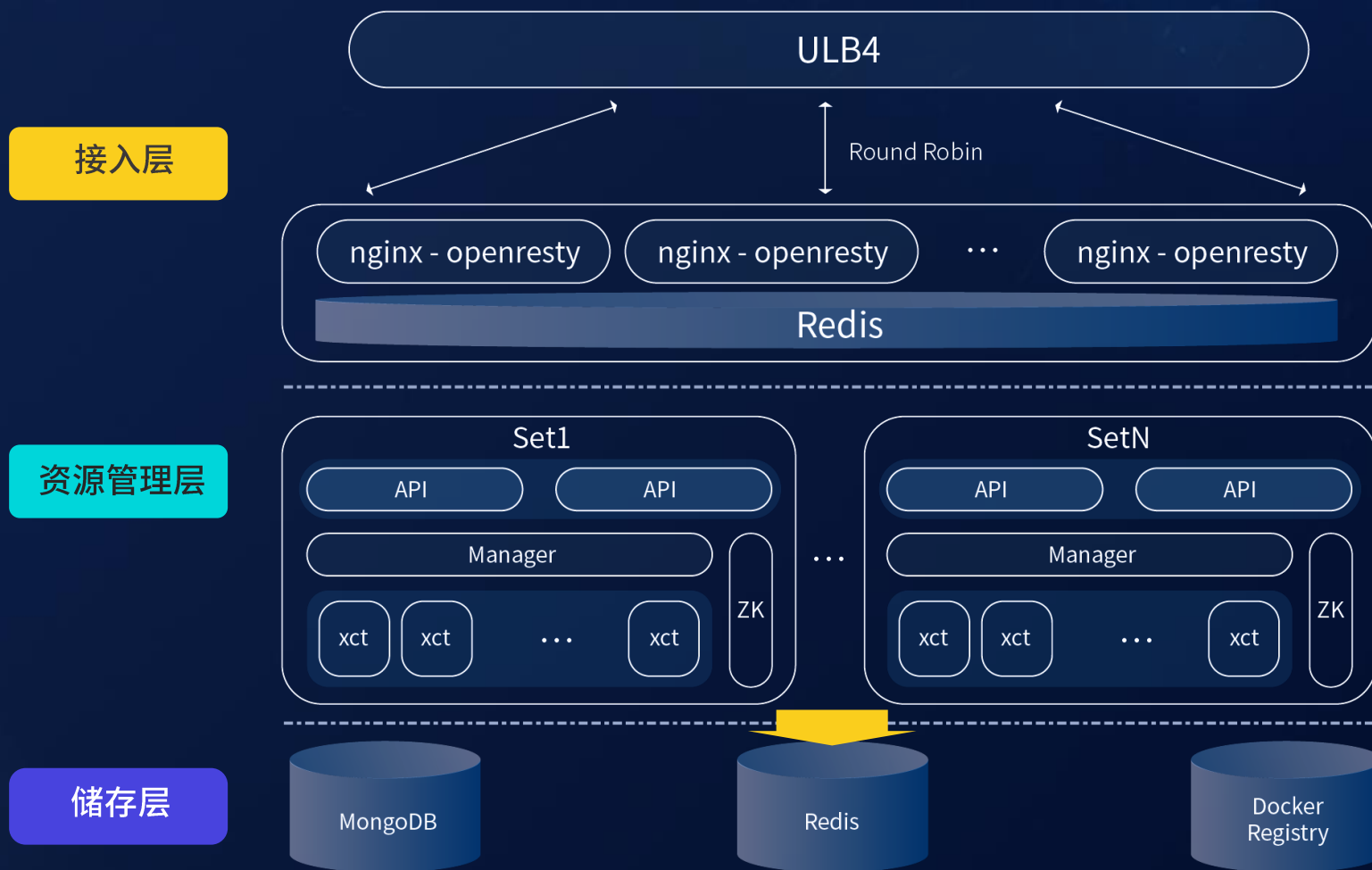
计算服务化

- 算法容器化
- HTTP API: 输入+算法+输出

III 计算平台



III 计算平台



III 计算平台

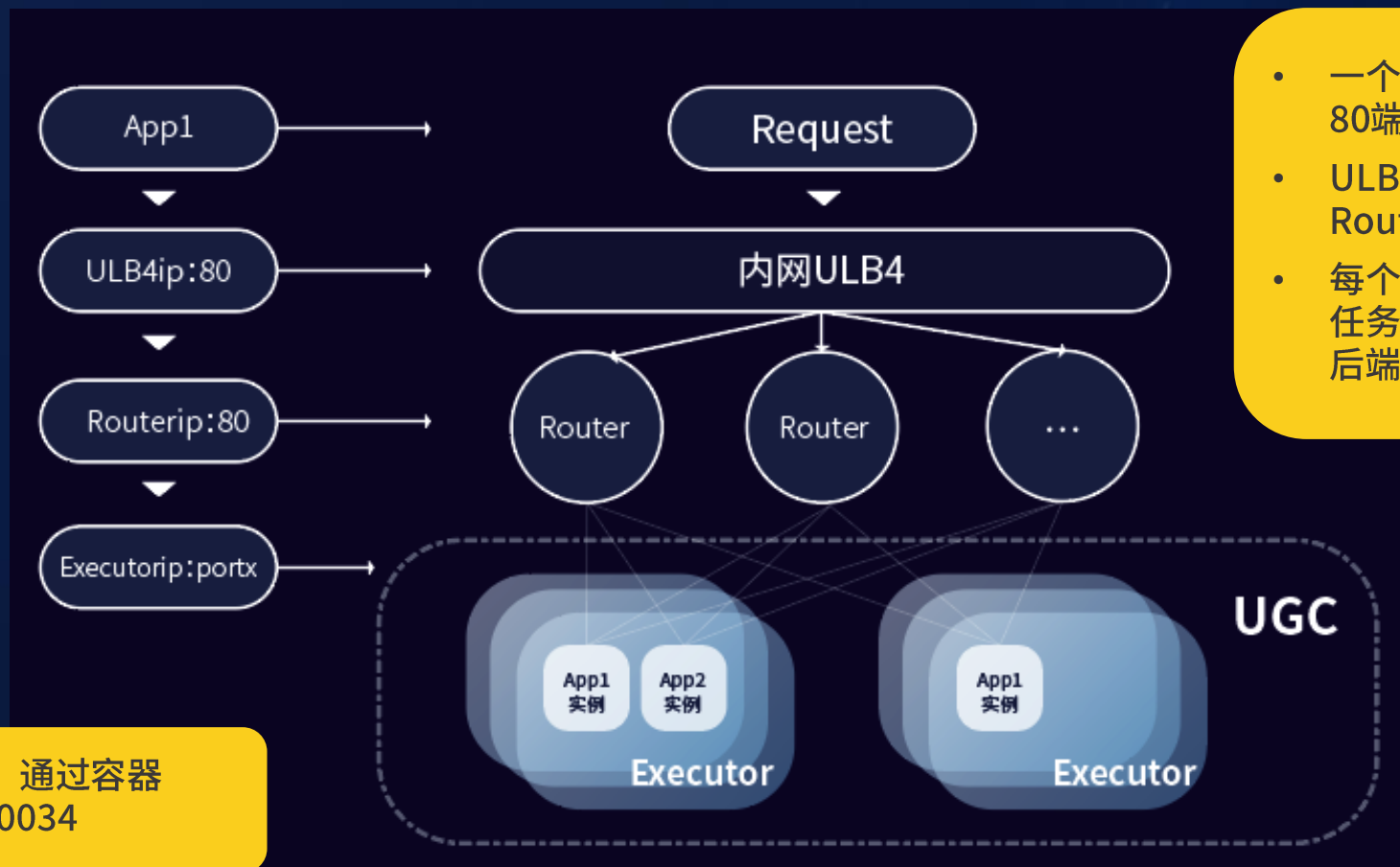
跨可用区容灾

按Set部署

灰度发布

所有模块可以平行扩展

App Engine



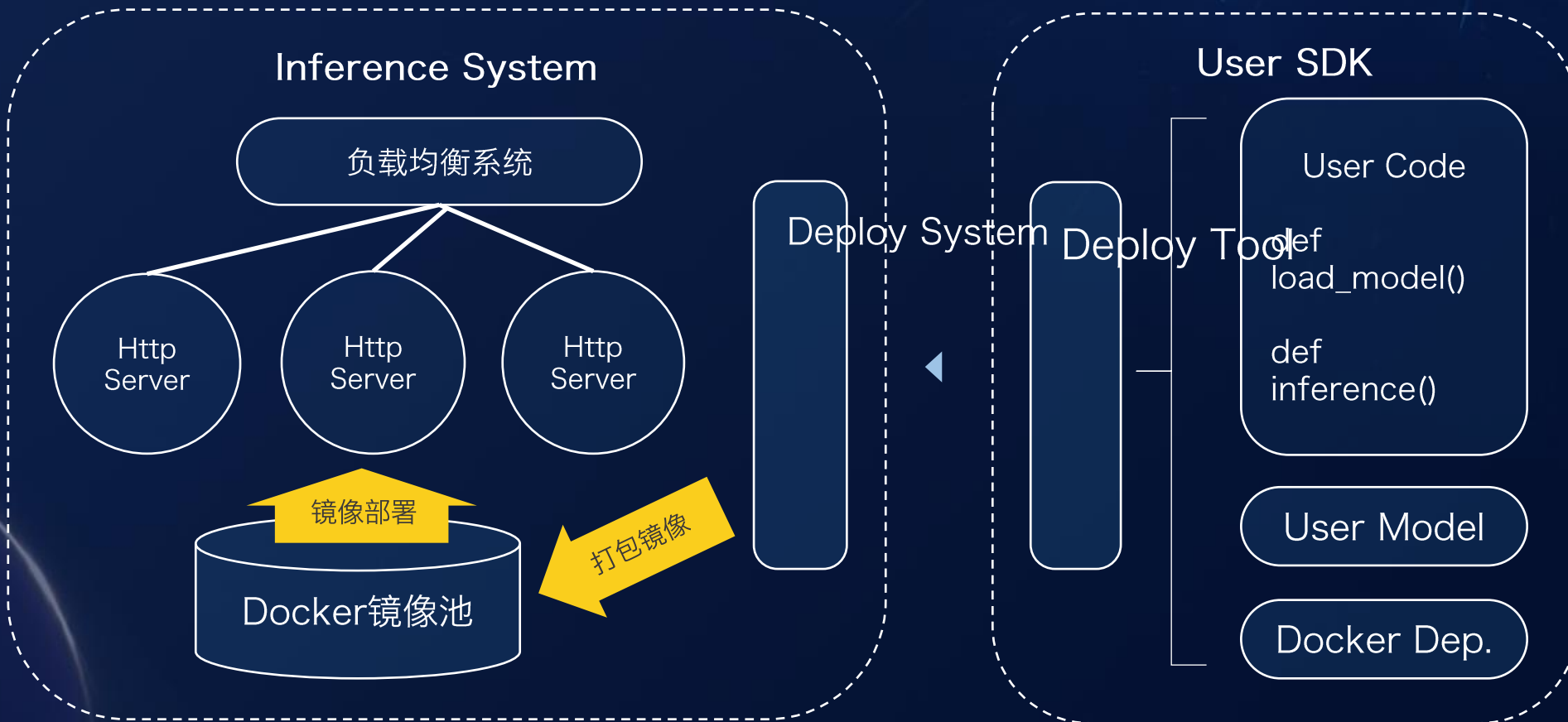
- 一个 App 访问 ULB 的 80 端口
- ULB 的端口与每个 Router 的端口一一对应
- 每个 Router 机器上提交任务给 UGC，UGC 负责后端的转发

用户App实例，通过容器
端口映射 80:10034

III 在线分类系统



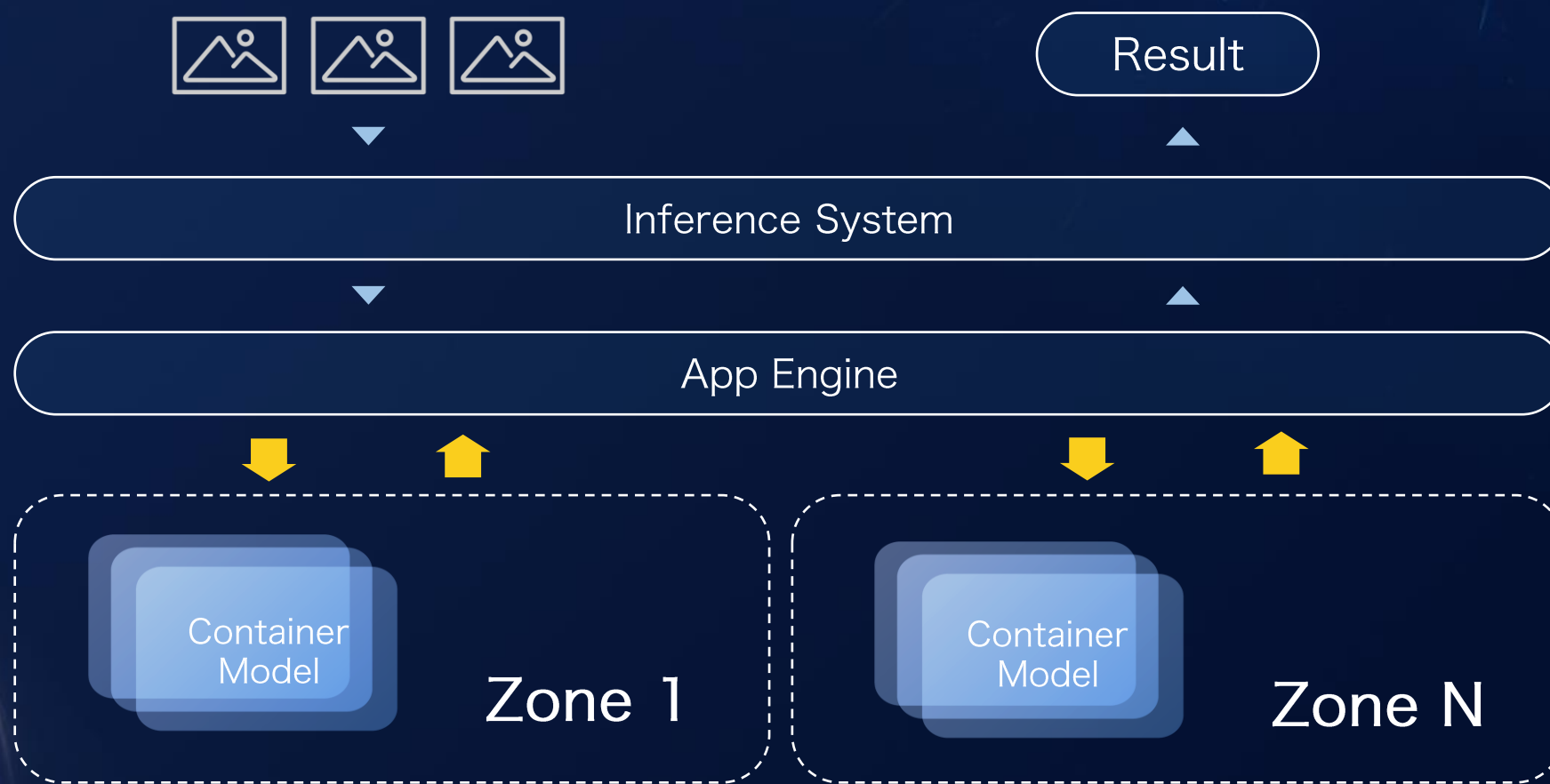
III 在线分类系统



05

应用案例

头像识别

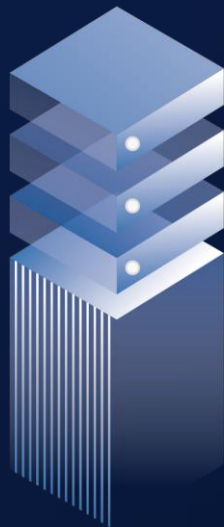


OCR

App



中控服务



Inference System

App Engine

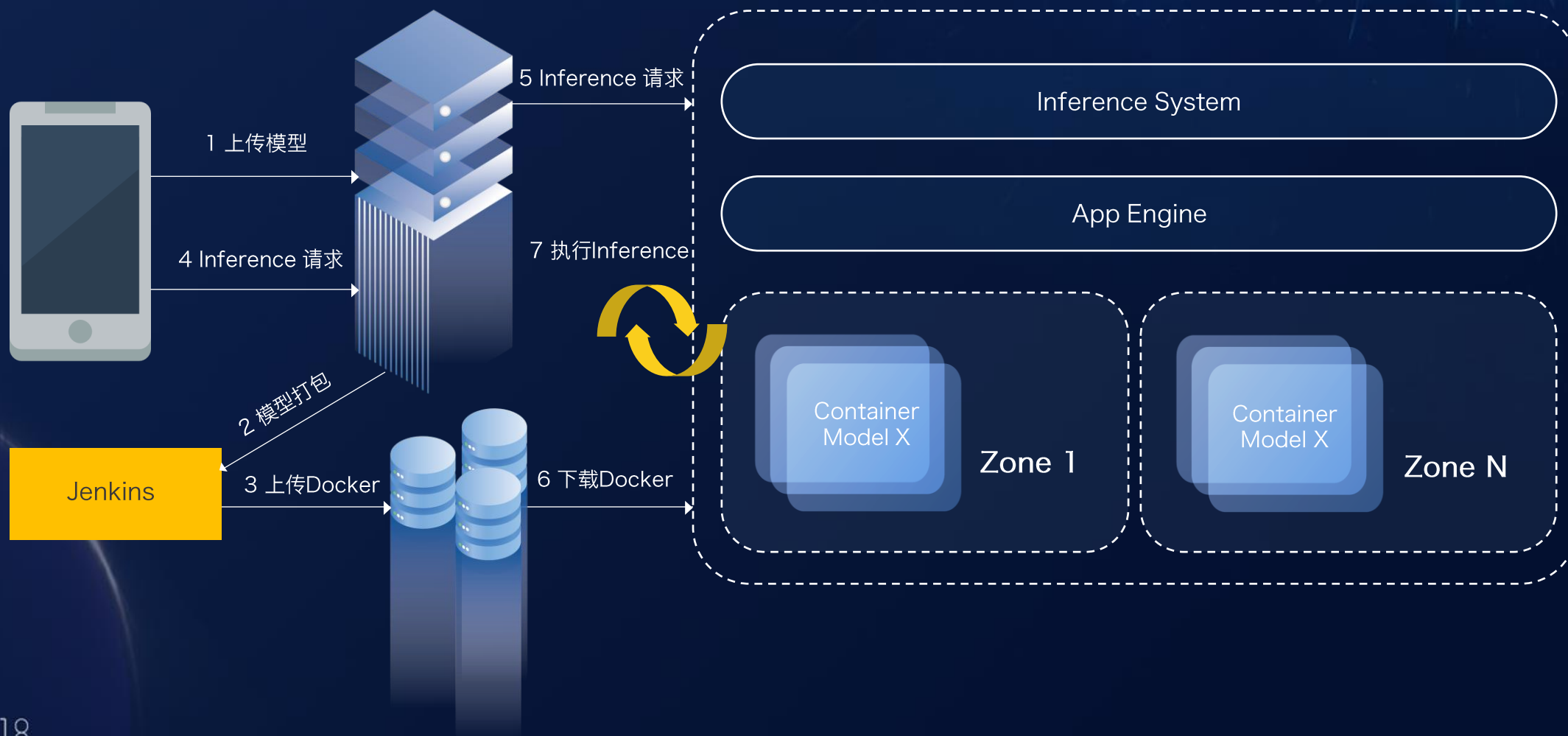
Container
OCR Model

Zone 1

Container
OCR Model

Zone N

AI应用市场



联系我们

UAI Inference UAI Train

Github:

<https://github.com/ucloud/uai-sdk/>

UAI Train:

<https://www.ucloud.cn/site/product/uaitrain.html>

UAI Service:

<https://www.ucloud.cn/site/product/uaiservice.html>

Contact: 4000188113

加入我们团队

Contact:

charlie.song@ucloud.cn

john.hu@ucloud.cn



THANKS