

# 云计算助力人工智能快 速搭建底层资源



徐强 互联网事业部解决方案架构师

# 目录

---

01

AI 落地的挑战

---

02

AI 落地解决思路

---

03

AI 落地解决方案

---

04

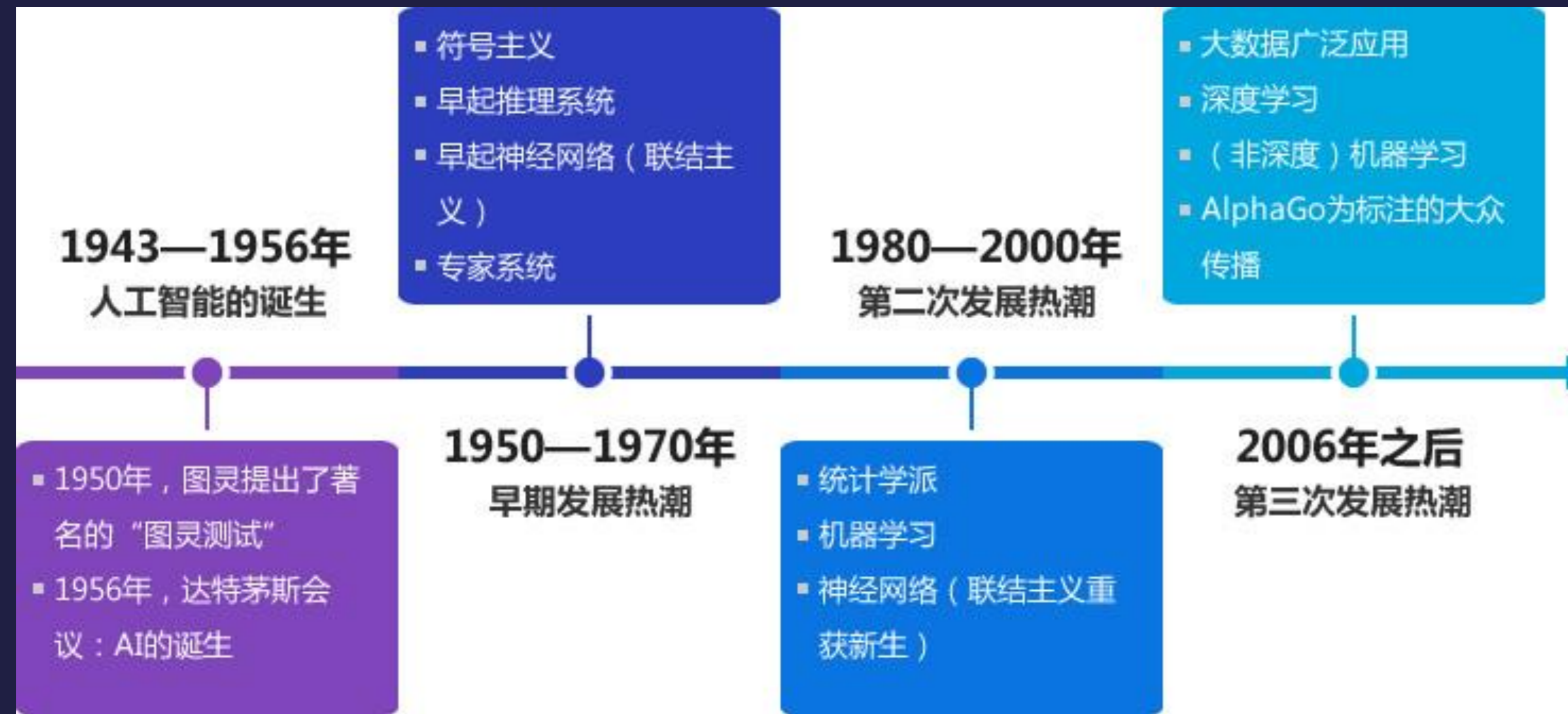
案例分享

01

# AI 落地的挑战

# 人工智能发展历史以及行业政策

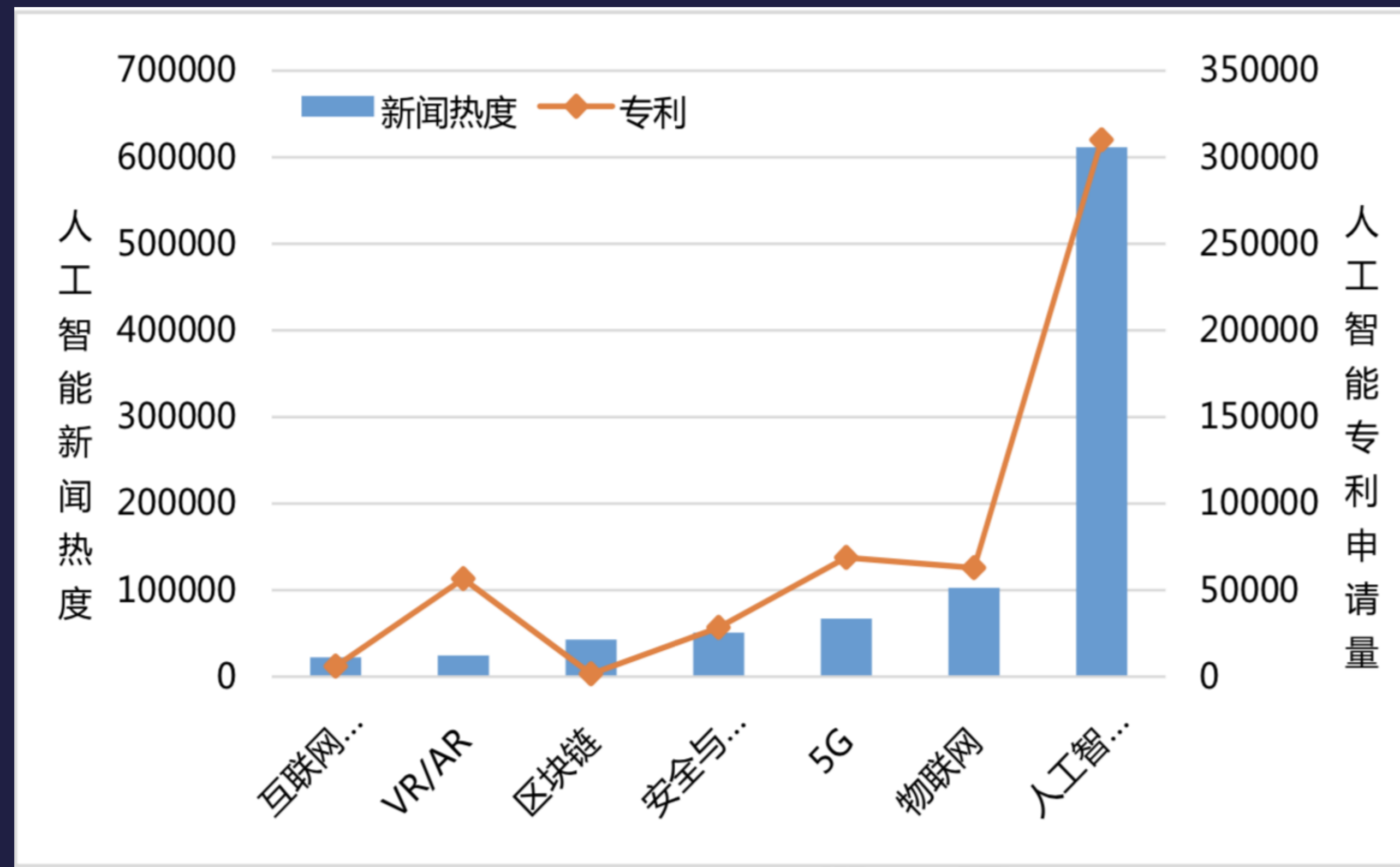
## 人工智能60年发展历程



### 我国将人工智能列为重点发展目标

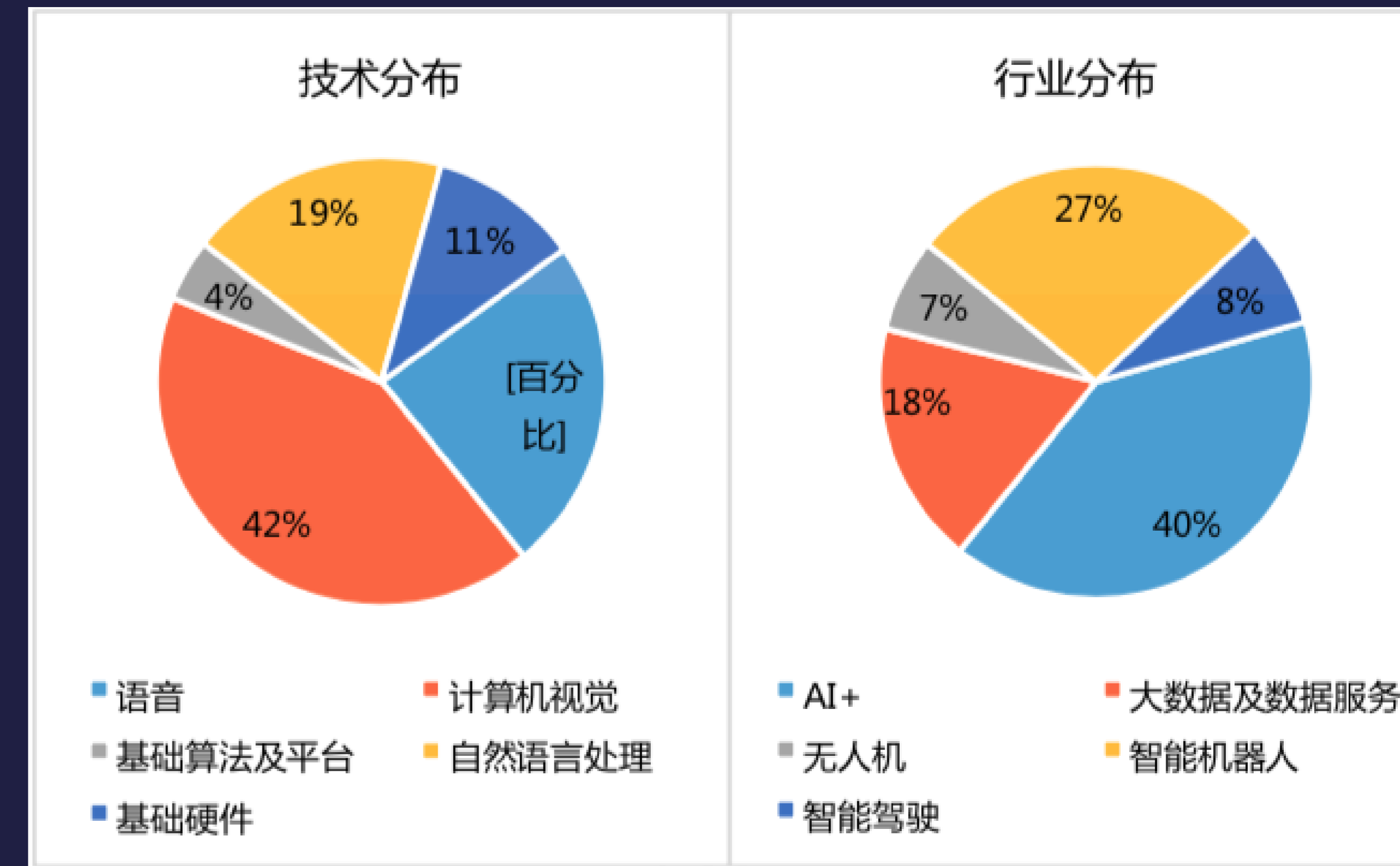
- 2015年6月：将人工智能列为重点发展目标
- 2016年3月：“十三五”提出突破人工智能技术
- 2016年4月：《机器人产业发展规划》提出重点研究人工智能
- 2016年5月：《“互联网+”人工智能三年行动实施方案》
- 2016年7月：《“十三五”国家科技创新规划》
- 2017年3月：《政府工作报告》首次写入人工智能
- 2017年7月：《新一代人工智能发展规划》出台
- 2018年9月：习近平致信祝贺世界人工智能大会开幕

# 人工智能成为行业焦点以及国内应用场景分



2017年全球信息通信热点领域新闻热度及专利申请数

人工智能 ( Artificial Intelligence ) 是全球2017年信息通信领域最热门焦点，远高于物联网、5G、安全与隐私、区块链等。



2017年国内人工智能企业应用场景分布

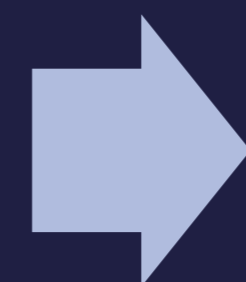
从技术分布来看：计算机视觉、语音、自然语言处理排名前三。  
从行业分布来看：AI+（融合医疗、金融、教育和安防等领域）、智能机器人、大数据及数据服务排名前三。

# 人工智能三部曲



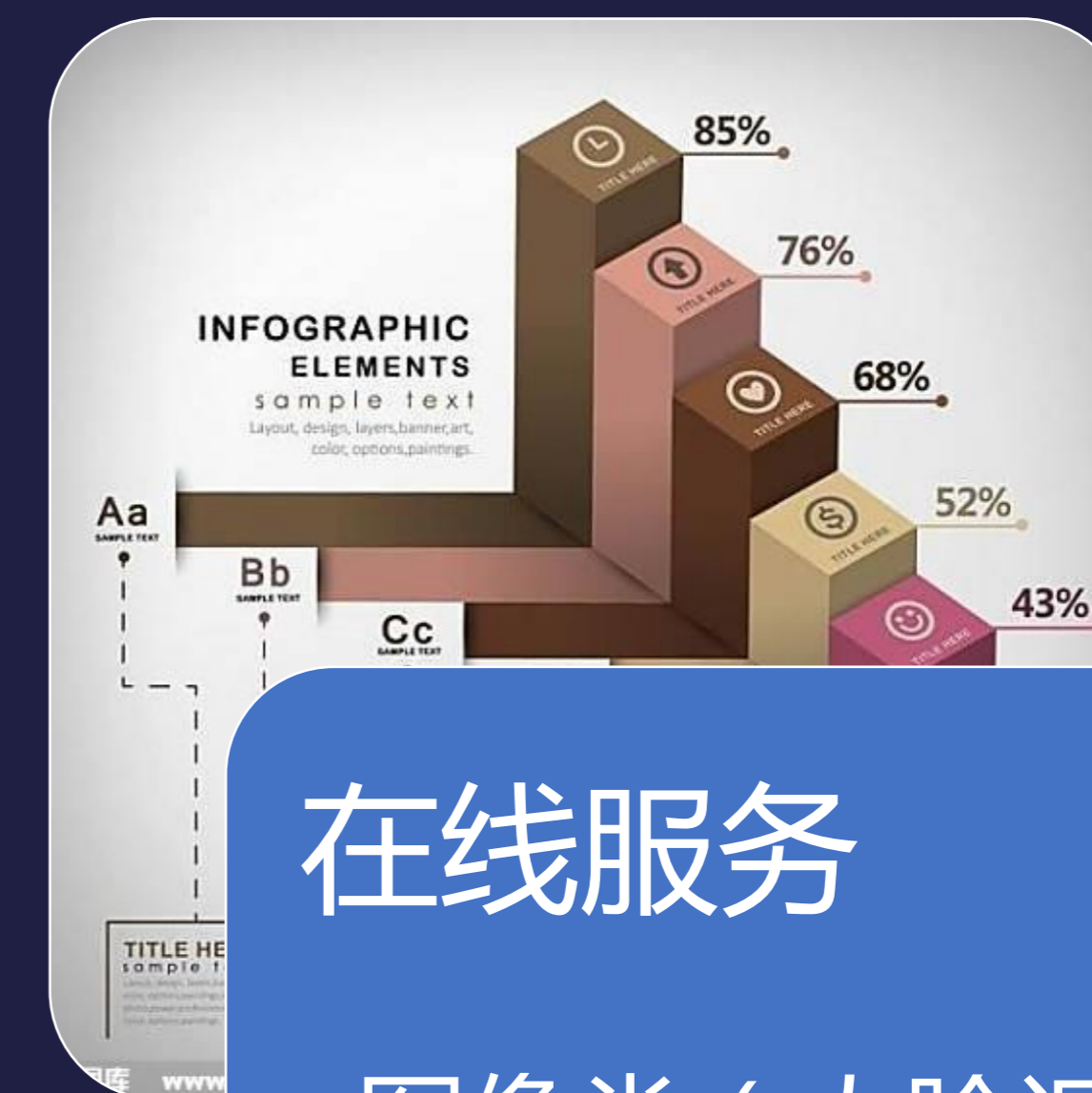
## 数据整理

- 数据收集
- 数据清洗
- 数据存储



## 模型训练

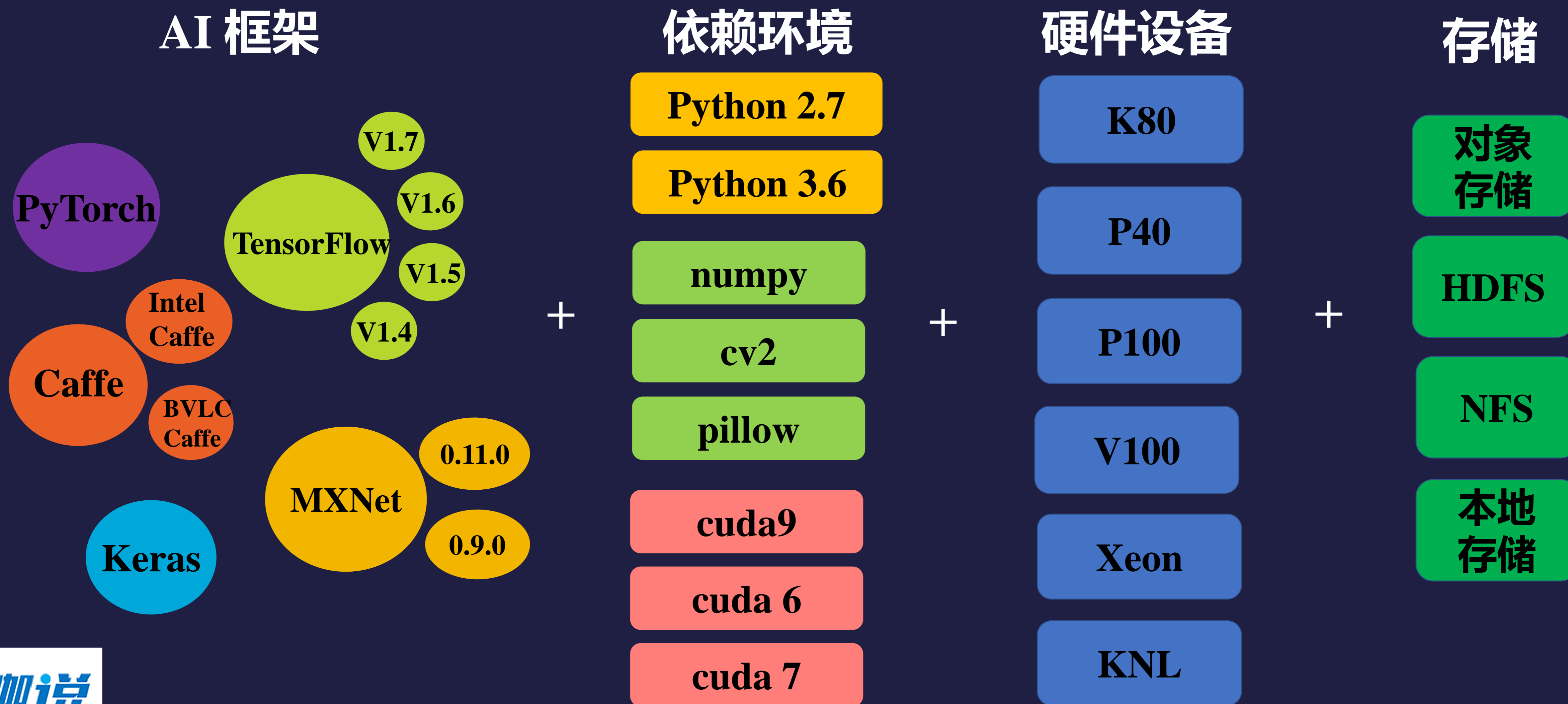
- 系统部署
- 框架/算法选择
- 模型训练



## 在线服务

- 图像类（人脸识别）
- 语音类（语音识别）
- 语言类（自然语言）

# AI的挑战I：基础环境



# AI的挑战II：AI 系统建设

## 算法兼容性

更好地兼容各类AI  
框架和算法

## 平台扩展性

平台具备扩展能力，  
支持业务规模的不断  
扩大

## 分布式化

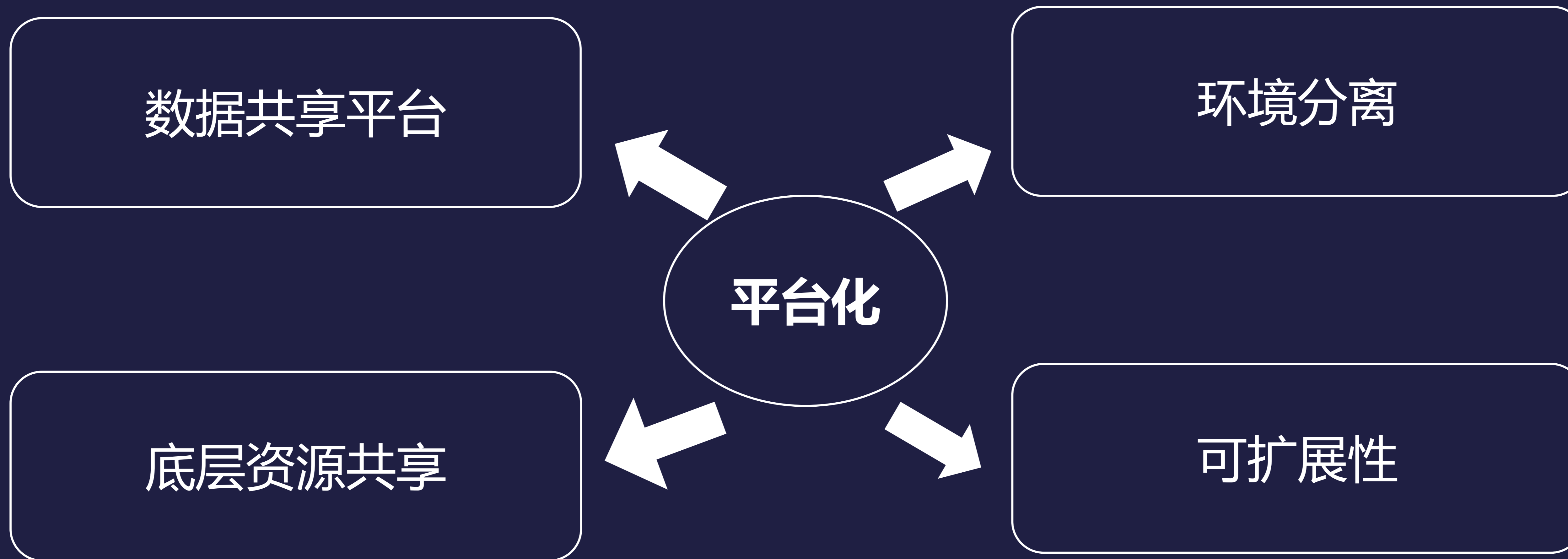
具备弹性伸缩的能  
力以及容灾能力



# 02

## AI 落地解决思路

# AI 落地挑战的解决思路



# 一、数据共享平台

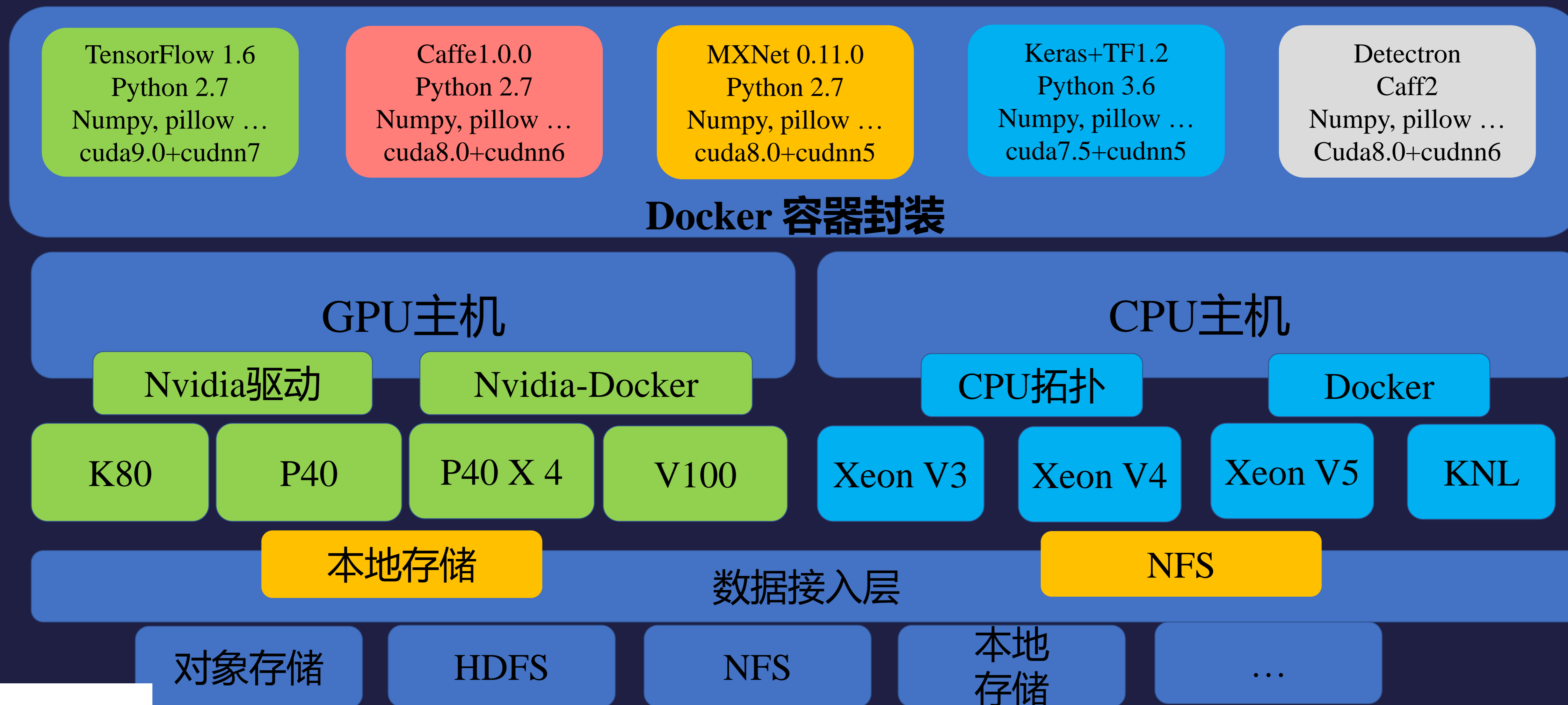
数据所有权不变的情况下，实现数据使用权的可信流通共享



# 安全屋引进的核心数据源

类型	客户名称	数据内容	
运营商	中国联通 中国电信	联通覆盖2.7亿活跃用户/电信覆盖2.4亿活跃用户	
		1.基础属性数据 2.话单数据	3.互联网日志数据 4.位置数据
	创数纪	电信全国数据、联通（江苏）	运营商基础数据、用户属性、行业标签、业务标签、场景标签
SDK	Mob	共8亿月活跃设备：	
		1.基础属性数据 2.LBS轨迹数据	3.APP安装列表 4.设备属性数据
垂直行业	闪云付	4,000万用户的理财和200万贷款信息数据	
	车轮网	月活过亿汽车人群用户	
	火眼云	3000万家企业的数据画像	
	大查柜	提供电商应用、每天产生约2000万条	

## 二、环境分离



## 二、环境分离：容器

隔离

运行环境完全隔离，不同任务之间不会产生软件冲突

预装

基础镜像内置各类基础软件环境，减少使用者环境准备开销

自由

可以自由安装各类软件包，封装各类算法

可重用

算法的容器镜像可以重复使用

### 三、海量资源共享

AI PaaS  
平台

AI Inference  
在线服务平台

AI Train  
训练平台

公有云  
基础架构

CPU

K80 GPU

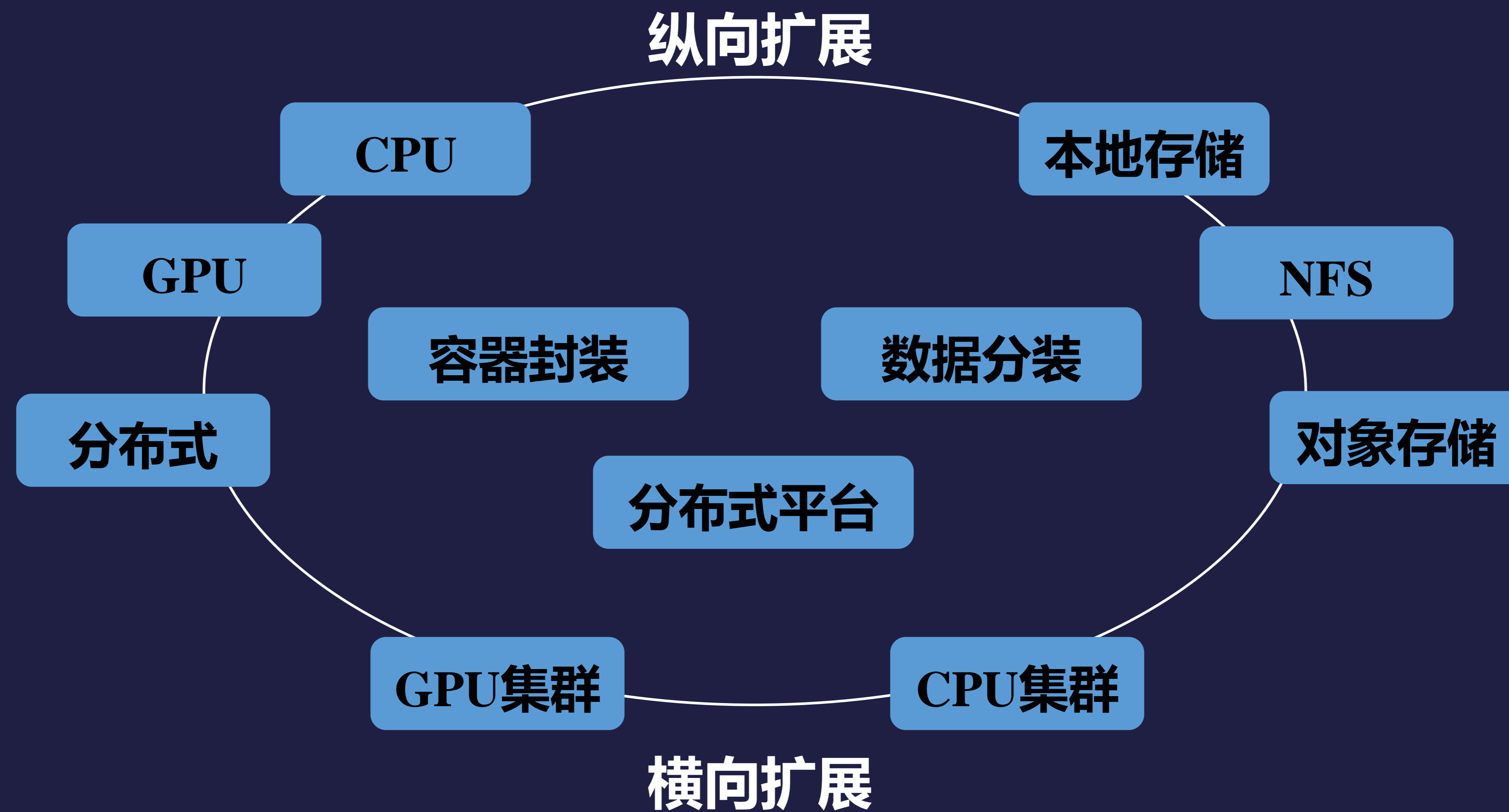
P40 GPU

P40X4 GPU

V100 GPU

海量计算资源

# 四、可扩展性

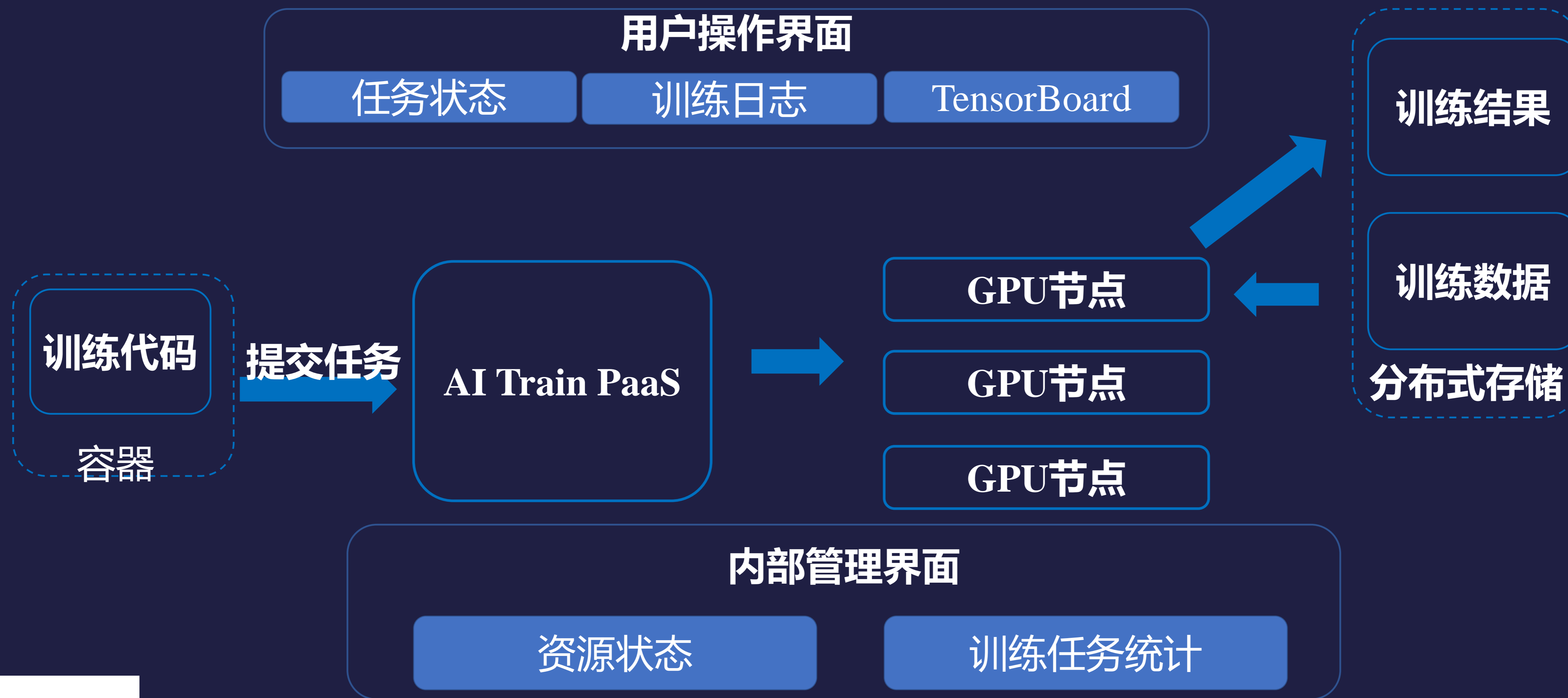




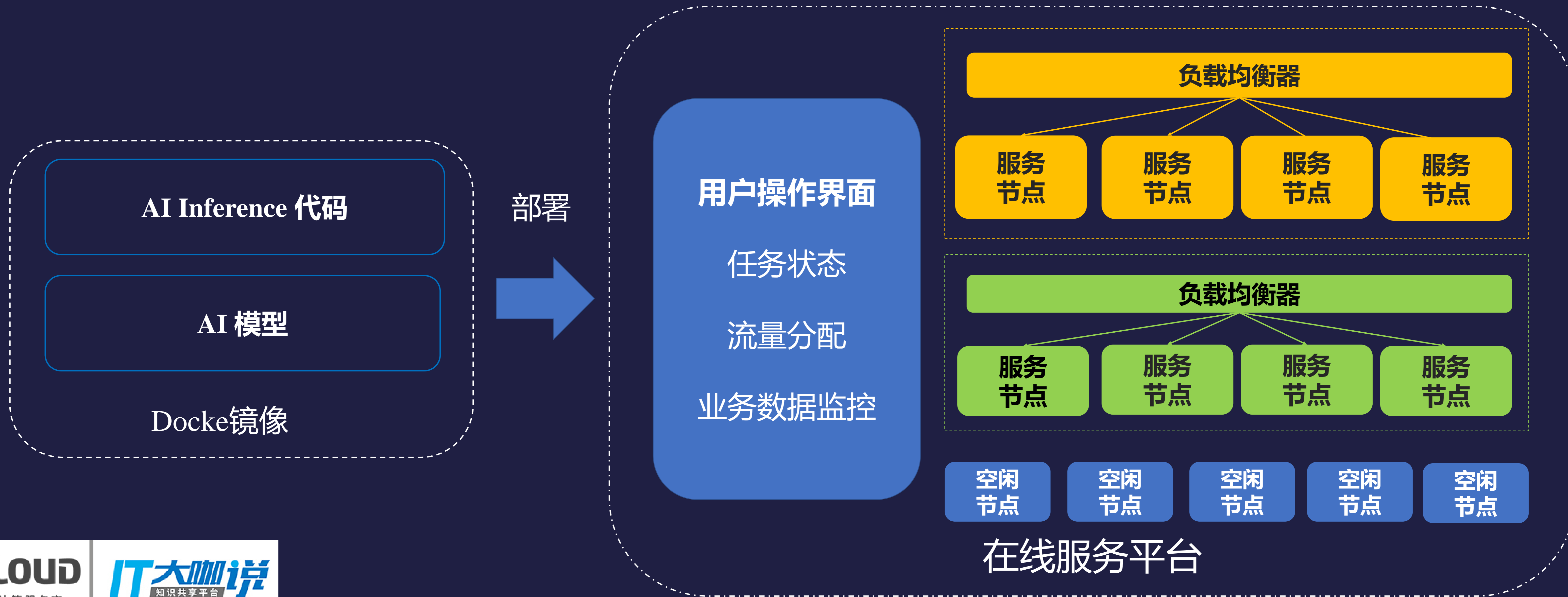
# 03

## AI 落地解决方案

# 一、分布式AI 训练平台 ( UAI - Train )



## 二、在线服务平台 ( UAI - Inference )



# 分布式 AI 在线服务平台 —— 高可用

负载均衡



# 公有云支持 AI 落地之 PaaS 服务

提供预置AI基础环境，包括NV GPU驱动、Cuda、TensorFlow/MXNet等框架



用户无提供充足CPU/GPU资源，可自由横向扩展  
需担心资源问题



环境封装

分布式

横向扩展

计费灵活



提供AI训练平台和AI在线服务平台，提供一站式AI服务  
用户无须自行搭建复杂的AI平台



基于秒级/分钟级的计费规则，按需收费  
用户无需担心资源浪费

# 公有云支持 AI 落地之 IaaS 服务

## 资源

- 充足的计算资源、存储资源、网络资源
- 降低AI研发过程资源采购、维护的成本

## 基础环境

- 提供虚拟机镜像、容器镜像等服务
- 降低AI研发、应用过程中AI环境部署的难度

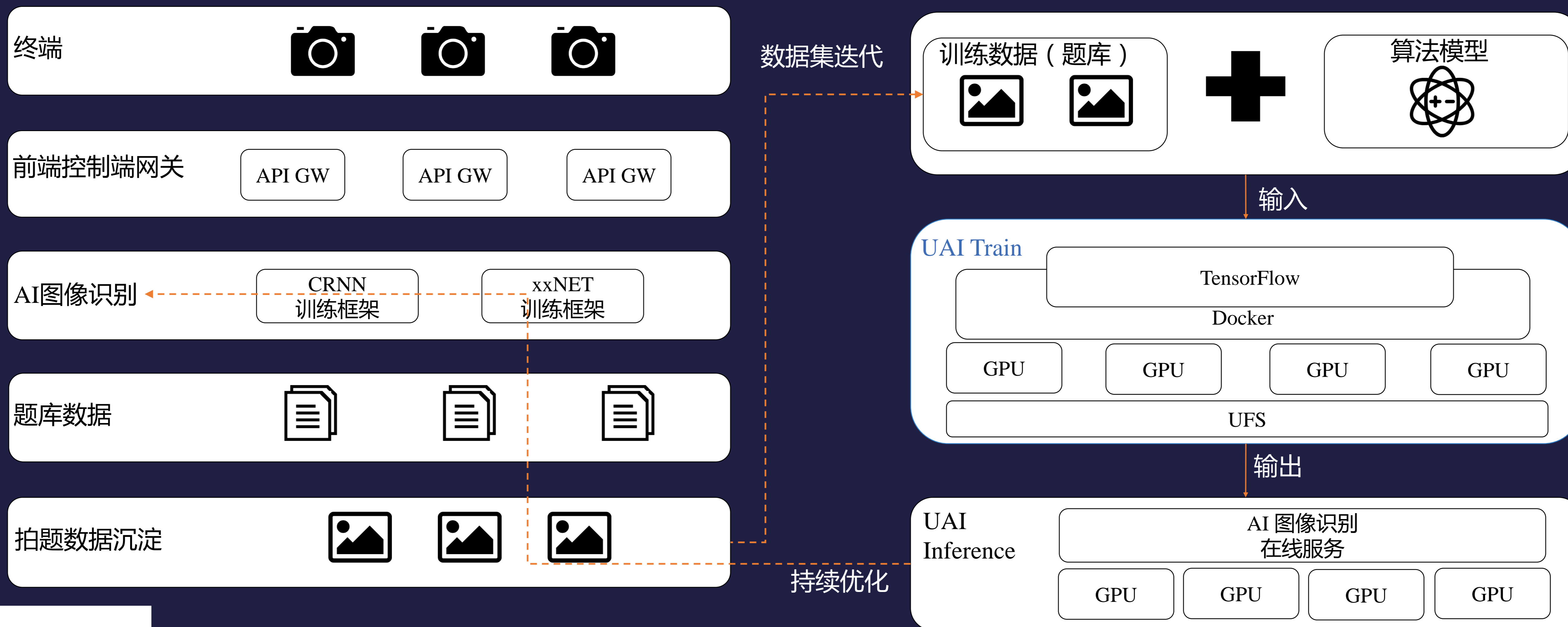
## 基础服务

- 提供诸如负载均衡(ULB)、分布式存储等基础服务
- 降低AI应用产品化过程的研究成本

# 04

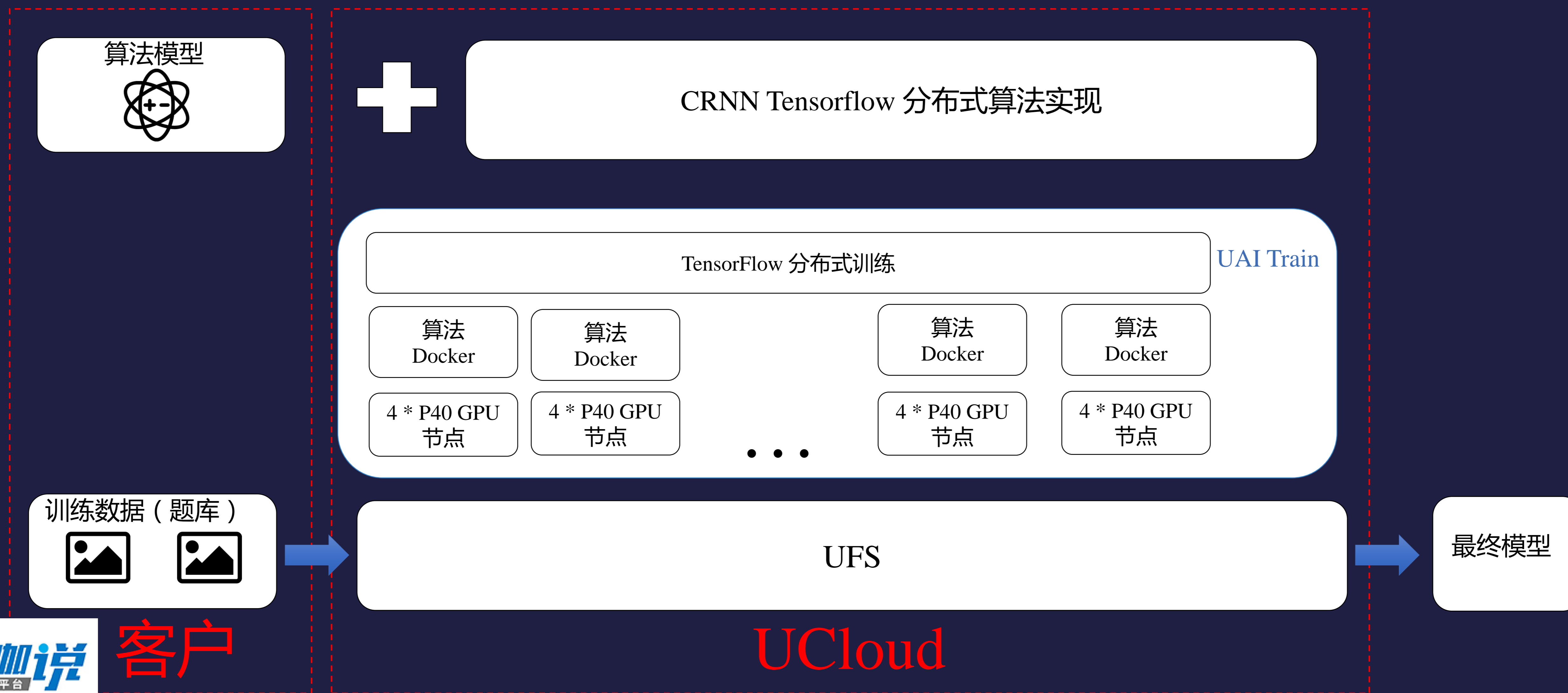
## 案例分享

# 案例分享一：作业盒子业务架构

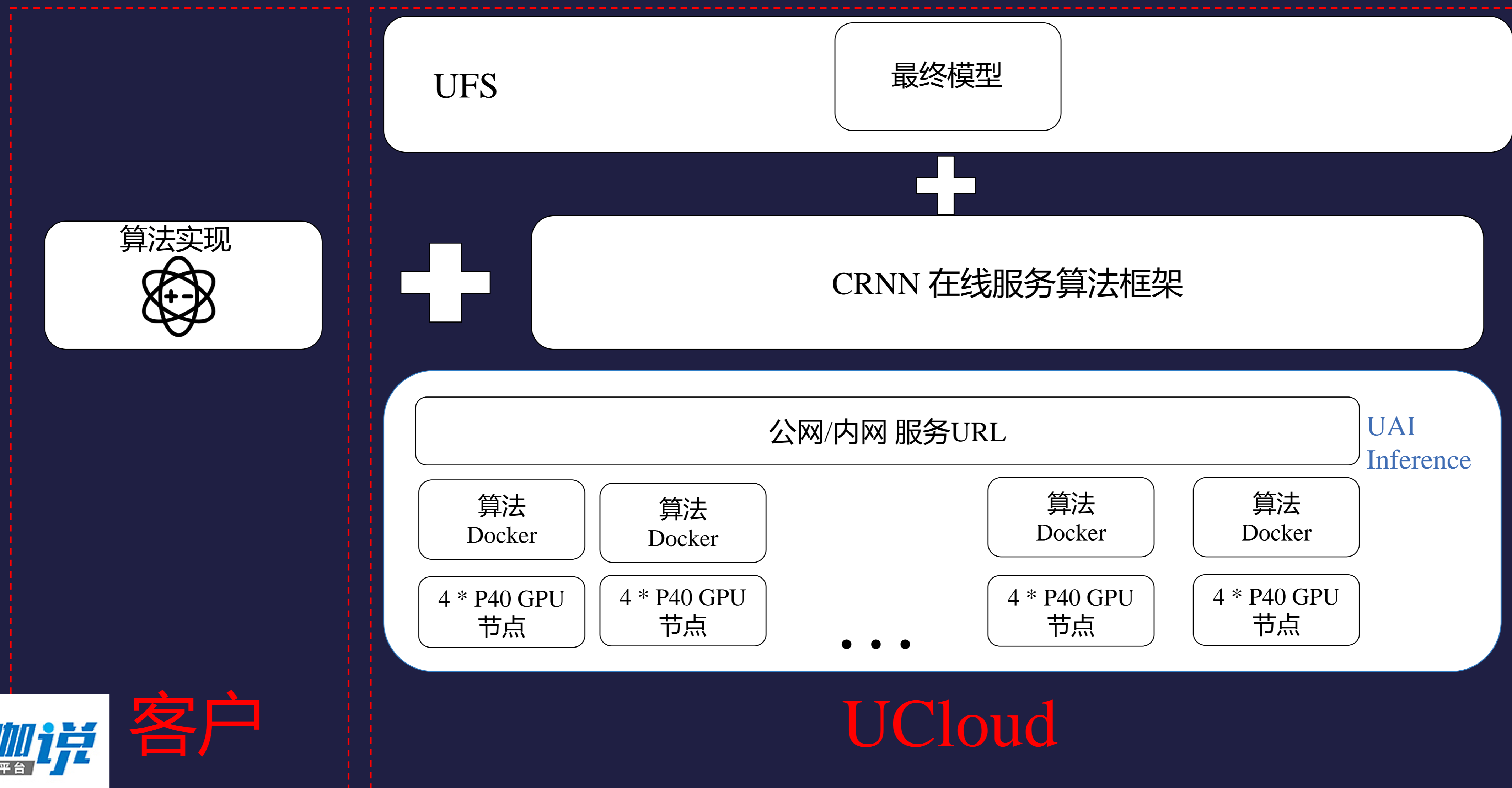




# 案例分享一：分布式训练作业题识别OCR



# 案例分享一：在线服务作业题识别OCR



# 案例分享二：AI Challenger全球挑战赛



安全屋

AI训练服务

AI在线服务

GPU云主机

AI Challenge全球AI挑战赛是由**创新工场**、**搜狗**、**今日头条**三家联合发起的全球性AI大赛。该大赛奖金总额超过200万人民币，设置了AI Challenger**视觉**和**翻译**两大类共五个赛道，自2017年9月开始，历时130多天，吸引了来自65个国家的超过一万名AI高手，各显身手、切磋技艺，并通过一次巅峰对决的现场答辩，最终决出获奖团队并分享奖金。

AI Challenge全球AI挑战赛在设立比赛规则同时，还通过UCloud AI Train训练平台为条件有限的参赛选手提供免费GPU资源支持，帮助他们圆梦AI。

大赛官网：<https://challenger.ai>

# 案例分享二：UCloud AI 全力助力选手



快

**Telsa P40显卡**

- 3840个CUDA核心
- 12 TFlops的单精度性能
- 47 TOPS的INT8性能

简

**环境自动部署**

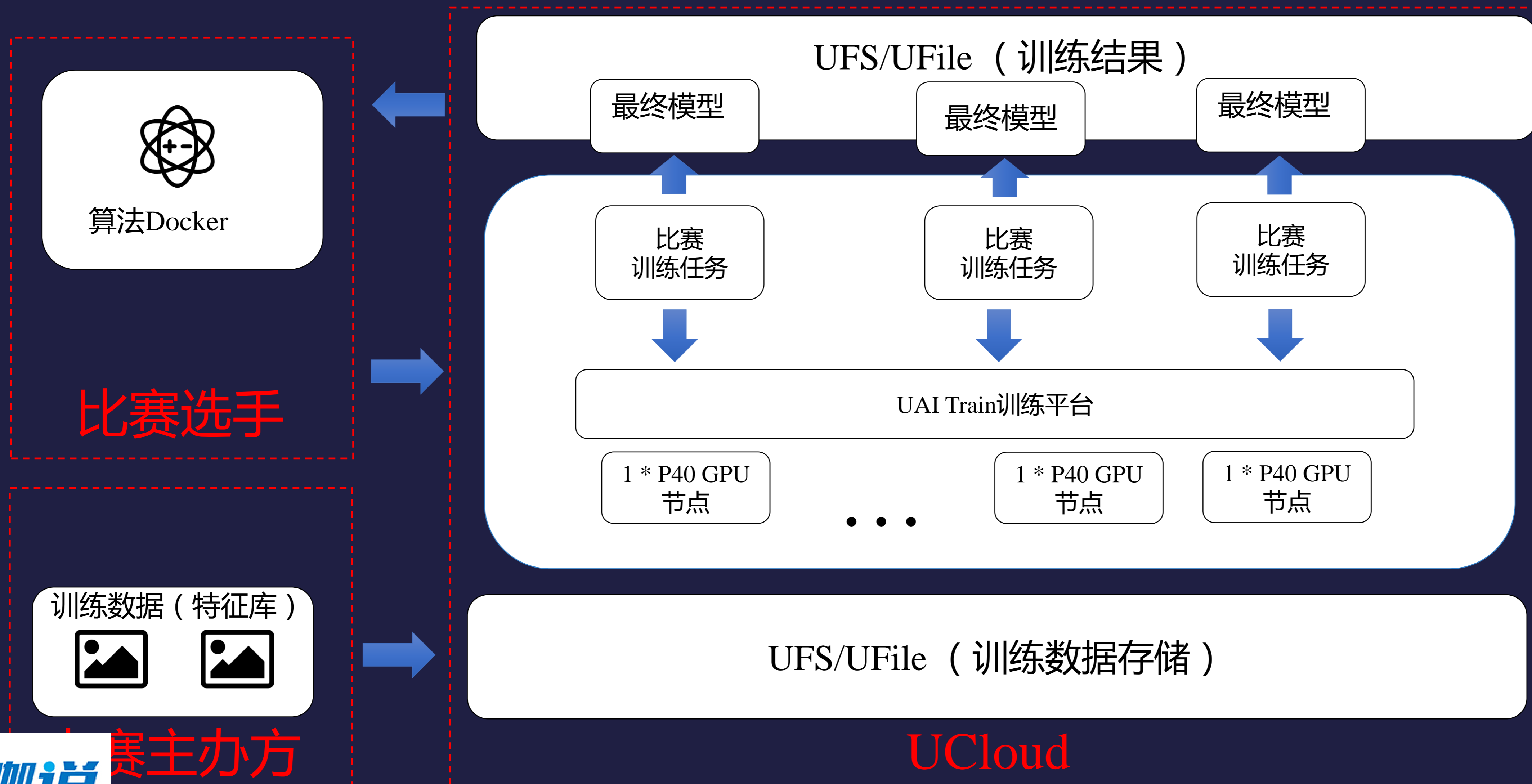
- 开箱即用基础Docker镜像
- 预装cuda, tensorflow
- 一键自动打包工具

省

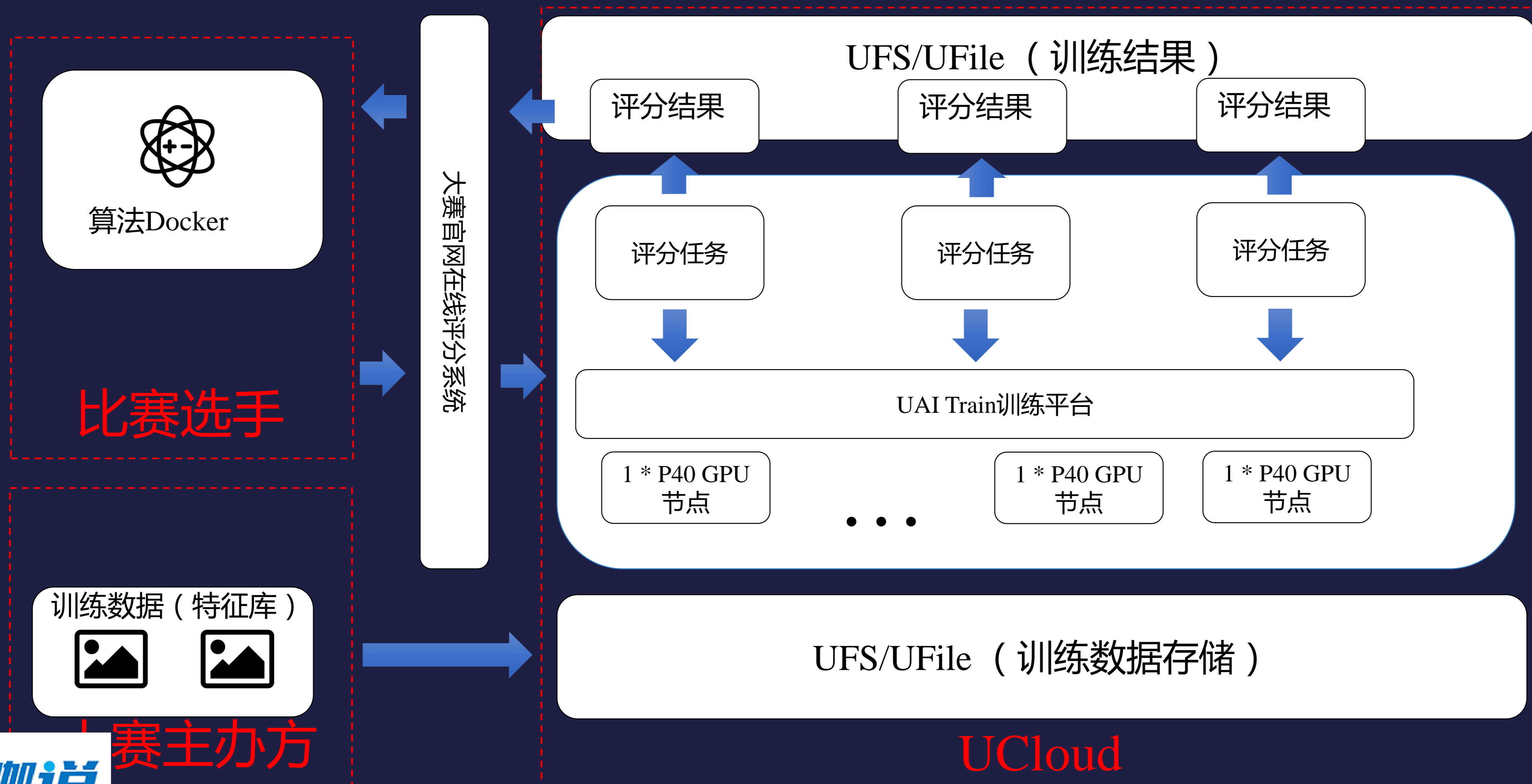
**按需收费**

- 按实际运行时间
- 精确到分钟
- 实时监控随时中止

# 案例分享二：AI Challenger训练资源



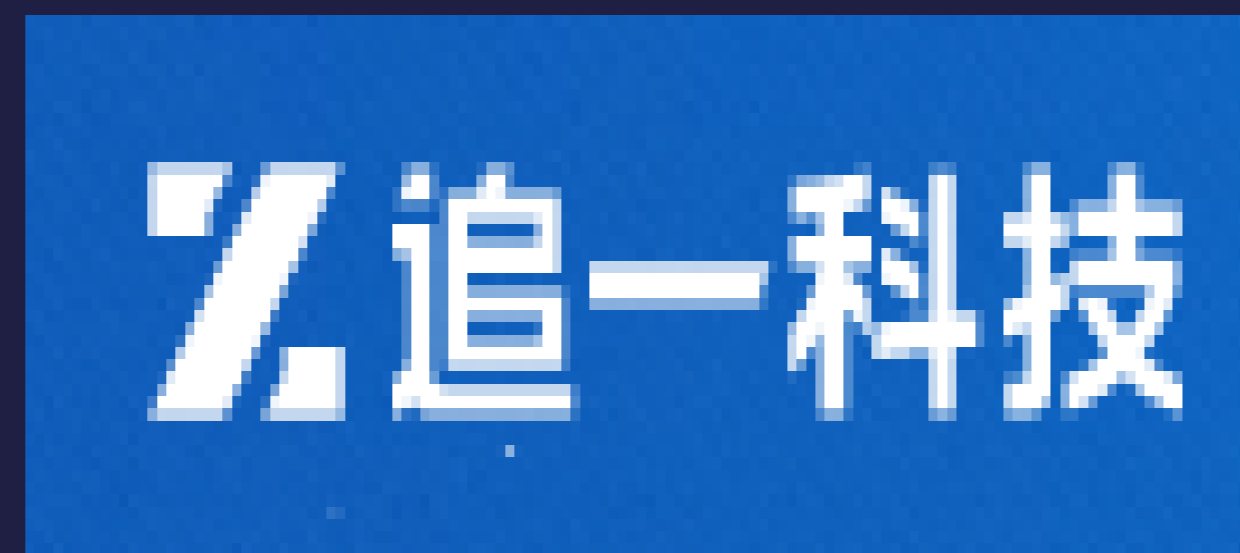
# 案例分享二：AI Challenger评分系统



# UCloud AI 客户案例



格灵深瞳



追一科技



Airdoc



作业盒子



远鉴科技



尺子科技



狗尾草



氦氦云

# THANKS

