

# Day 18 华为云深度学习技术起底

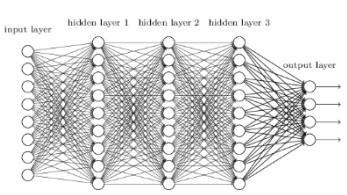


# 深度学习/神经网络令人兴奋的突破和趋势


技术发展

应用突破

2006



2013



95%



99%

2014~2016



2016~至今



2011



Siri

2012

Google's self-driving car



2014



amazon echo

2016



AlphaGo



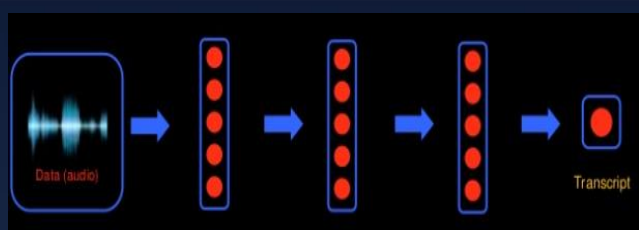
amazon go



# 深度学习在很多领域取得了历史性的突破

被认为已经解决或接近解决的问题

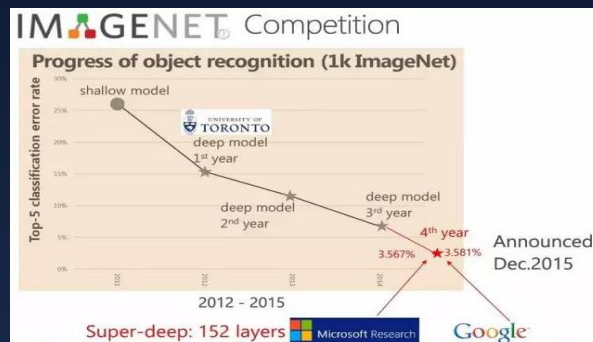
## 语音识别



业界最好水平：

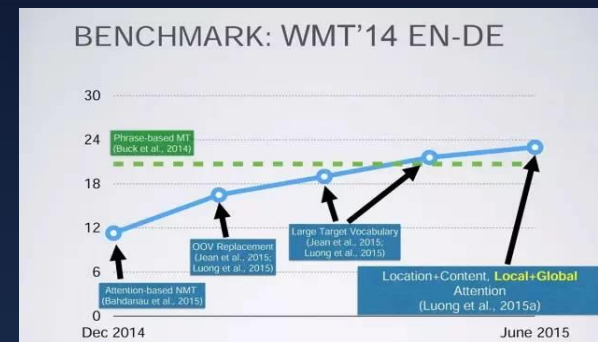
CER 3-5%, WER 9-11%，均已达到甚至超过人的平均能力

## 静态图像识别



在一些场景限定领域已经超越人类

## 翻译

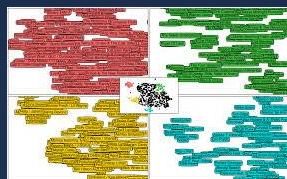


准确度和自然度上还没有超越人类

正在解决的问题（非完全列表）



自然语言对话



用户画像



行为建模控制



搜索和推荐



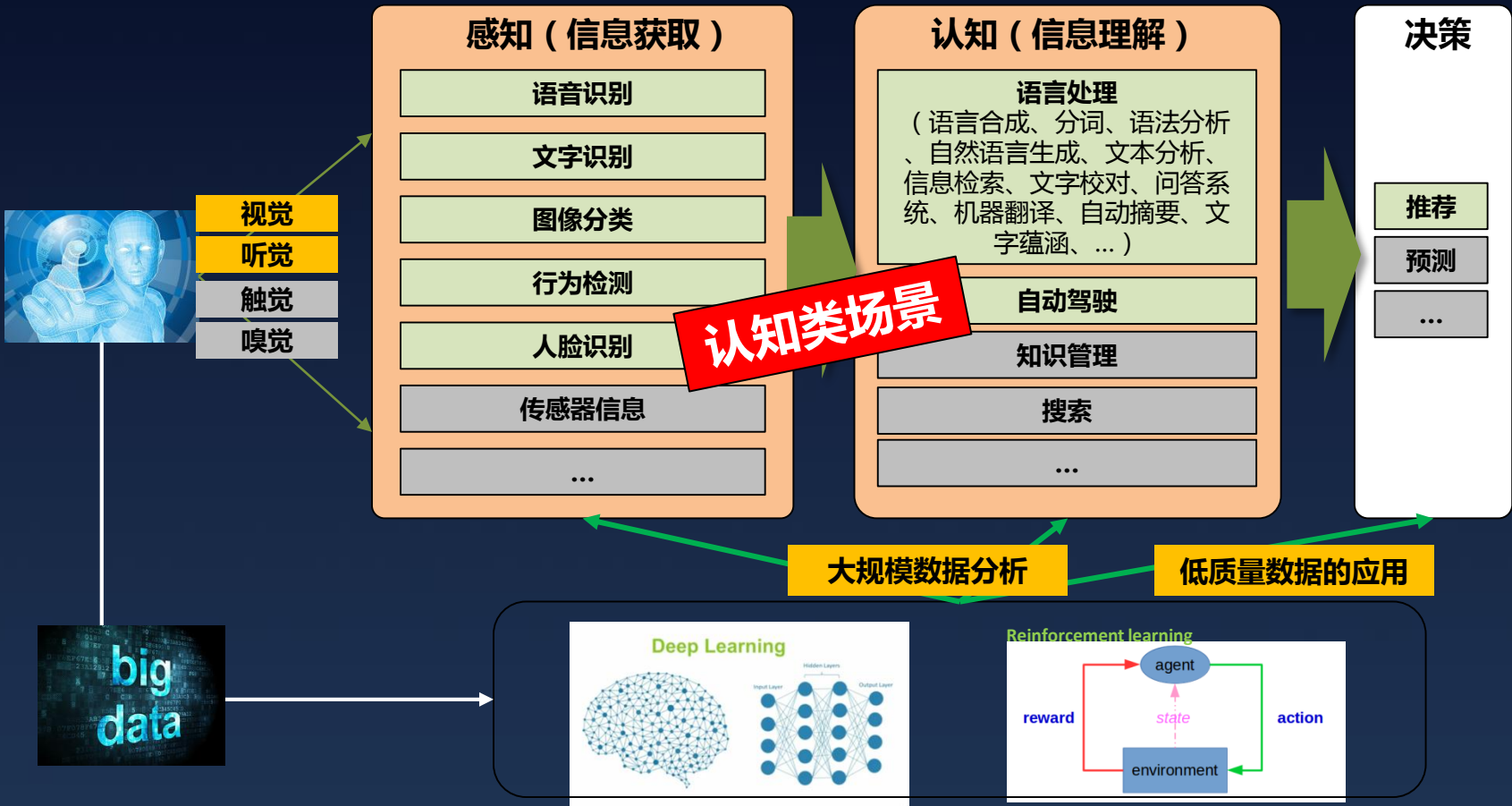
图像语义



个人健康

# 应用场景：感知认知类（视觉/语音/NLP）的业务场景为主

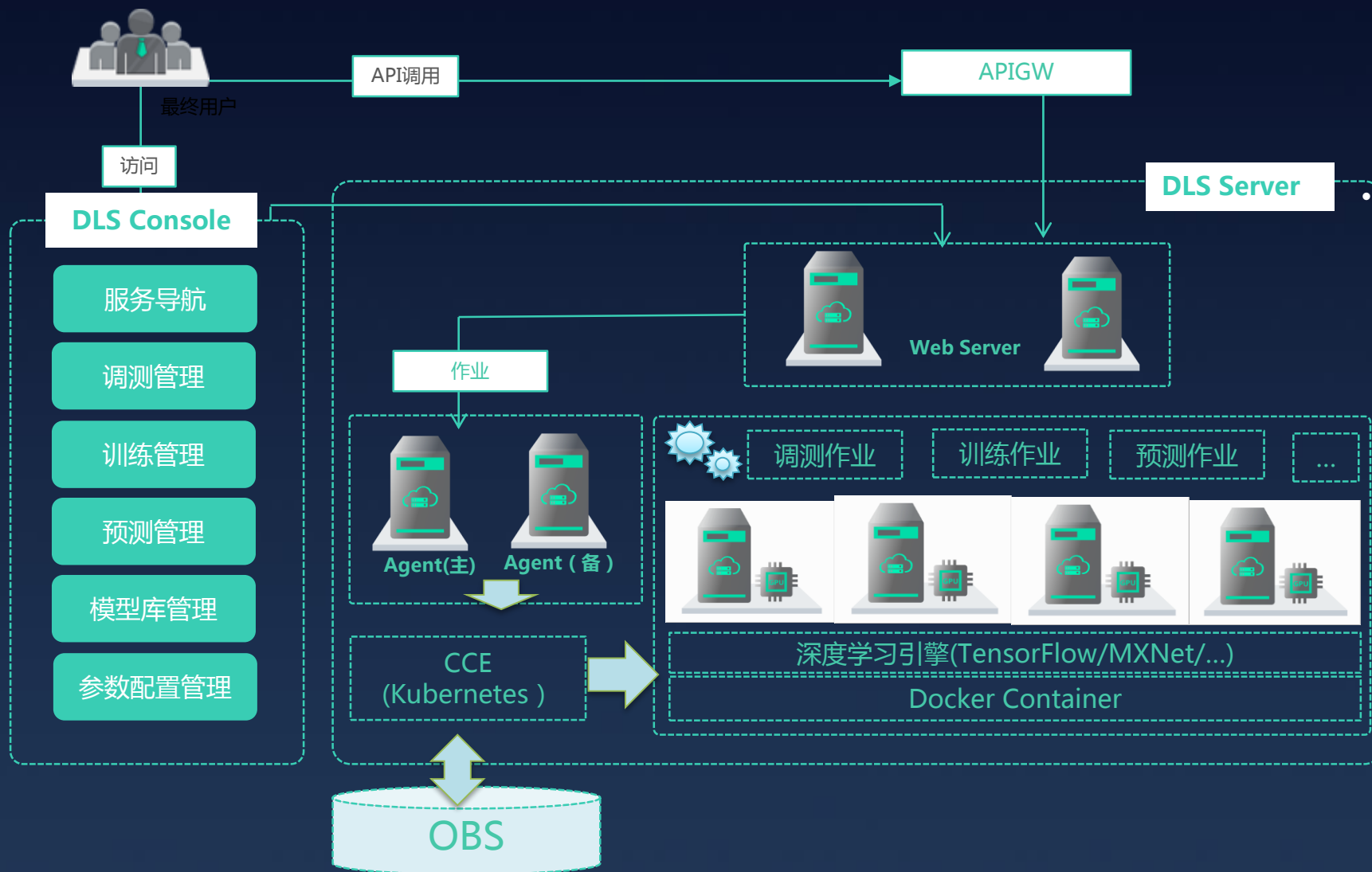
机器模拟人的行为的几个过程



过去3-4年，深度学习在**计算机视觉**，**语言处理**，**自然语言领域**取得了较大的突破，并且其技术推广和应用速度超过人们的预期。自2014年以来，100家AI创业公司就有9亿1千万美元的多轮融资，技术上面会大量的使用深度学习，计算机视觉、自然语言、语言占比超过40%



# 华为深度学习服务介绍



DLS ( Deep Learning Service ) 是基于华为云强大高性能计算提供的一站式深度学习平台服务，内置大量优化的网络模型算法，以兼容、便捷、高效的品质帮助用户轻松使用深度学习技术，通过灵活调度按需服务化方式提供模型训练、评估与预测。

# 华为深度学习服务介绍

相比于自建私有集群，使用华为深度学习服务的优势为：

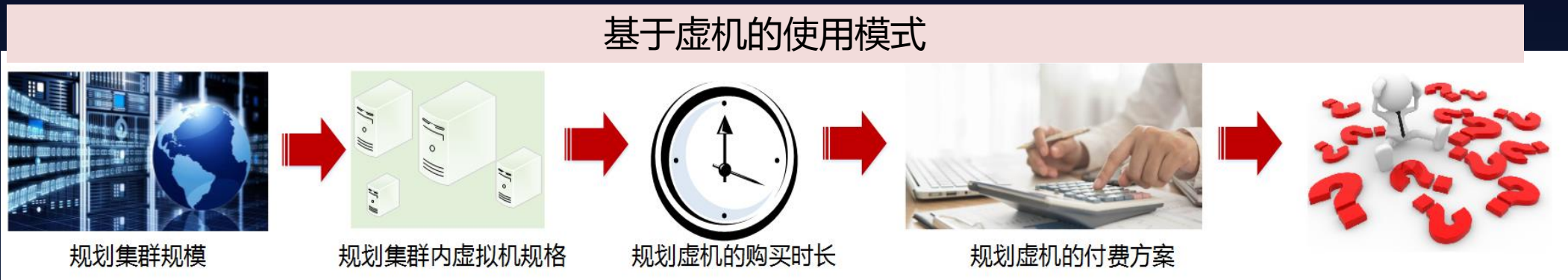
## 关键竞争力

- 端到端一站式平台, 企业级平台，让使用者专注业务应用，无需感知集群；
- 训练/预测作业实例**按需、按秒计费**，降低用户投入成本；
- 大规模集群，分布式训练加速比**>0.8**，训练从周级降低到小时级
- 打通数据快递服务/数据标注服务，给用户物美价廉的体验
- 预置高效的领域模型库，方便用户使用
- 调度与引擎解耦，新的AI框架可以被快速集成

关键项	自建私有集群	华为深度学习云服务
一站式调测/训练/推理UI易用界面	需自研	支持
费用	高	秒级按需计费，作业真正运行时才收费
GPU	由于GPU卡更新换很快，私有集群难以迅速体验更新的GPU卡，自购成本高	可以使用上最新的GPU，如V100等
分布式优化器增强	不支持	支持（批数据并行、EASGD、MA、LR自适应等多种分布式SGD）
模型库	分类模型、传统ML模型等	分类模型、NLP模型、检测模型、OCR模型等
分布式训练性能	慢 ( 5个以上节点训练收敛速度就开始明显下降 )	快 ( 100GPU规模分布式线性度0.8以上 )
分布式训练难易程度	难（单机转分布式实现，难）	简单（单机/分布式同一套代码，算法人员无需关注分布式）
是否需要运维人员	需要，成本高	不需要

# 关键价值特性一：按需使用，成本低

按照VM + GPU  
的资源粒度计费



按业务实际所需  
的计算资源计费

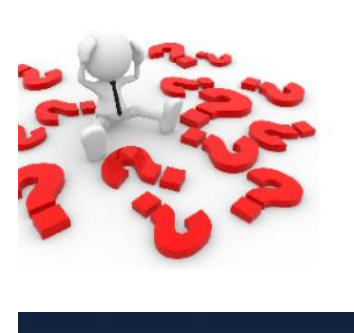

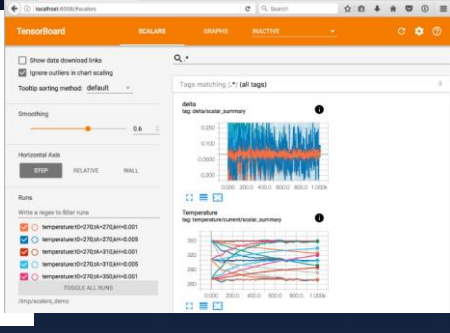

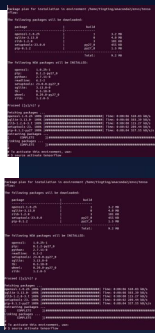

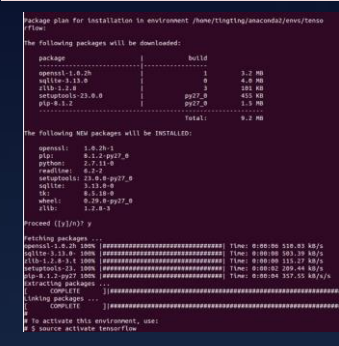


- 按照业务实际所需来分配资源
- 更准确、更灵活、更低成本
- 提供 1000+ GPU节点规模训练，GPU训练加速比0.8+

# 关键价值特性二：端到端易用的操作平台

基于开源，在shell命令操作和调试

基于开源



基于UI的易用操作平台

按业务实际所需的计算资源计费





深度学习服务

- 总览
- 预置模型库
- 开发环境管理
- 训练作业管理
- 预测作业管理
- 作业参数管理



1 代码开发与调测

用户可在“开发环境管理”页面，创建开发环境进行代码的开发与调测。或者在“预置模型库”页面，使用预置的模型，无需自行编写模型训练代码。



2 训练模型

用户可在“训练管理”页面，创建训练作业，进行模型训练。



3 预测服务

1. 用户可在“预测管理”界面，使用已训练好的模型创建预测作业，部署预测服务； 2. 用户下载并修改client样例代码工程发起预测请求，获取预测结果。




- 开发调测 训练 推理，实现端到端体验；
- TensorBoard、Jupyter Notebook、Timeline



# 关键价值特性三：超参自动选择、模型自动设计

为应用开发者，提供丰富的预置模型库



Deep Learning Service

Overview

[Pretrained Model Repo](#)

DevEnviron Mgmt

Training Job Mgmt

Visualization Mgmt

Inference Job Mgmt

Job Parameter Mgmt

Pretrained Model Repo

Model Name	Use of Model	Engine Type	Model Precision
SegNet_VGG_BN_16	--	MXNet, MXNet-1.1.0-python2.7	89%(pixel acc)
Faster_RCNN_ResNet...	Object detec...	MXNet, MXNet-1.1.0-python2.7	80.05%(mAP)
Faster_RCNN_ResNet...	Object detec...	TensorFlow, TF-1.4.0-python2.7	73.6%(mAP)
ResNet_v2_50	Image classi...	MXNet, MXNet-1.1.0-python2.7	75.55%(top1), 92.6%(top5)
VGG_16	Image classi...	TensorFlow, TF-1.4.0-python2.7	70.5%(top1), 89.7%(top5)
ResNet_v1_50	Image classi...	TensorFlow, TF-1.4.0-python2.7	74.2%(top1), 91.7%(top5)

Training Data Set

Key Running Parameter

train\_url = s3://

model\_name = resnet\_v1\_50

checkpoint\_url = s3://bxi/0001/resnet\_v1\_50/

+ Add Running Parameter.

\* Computing Node Specifications

32Core(s)   256GiB   4*P100	16Core(s)   128GiB   1*P100	56Core(s)   512GiB   1*P100   4TB(no)
56Core(s)   512GiB   1*P100   4TB	8Core(s)   32GiB	

\* Computing Nodes

- 1 +

Price Free in OBT

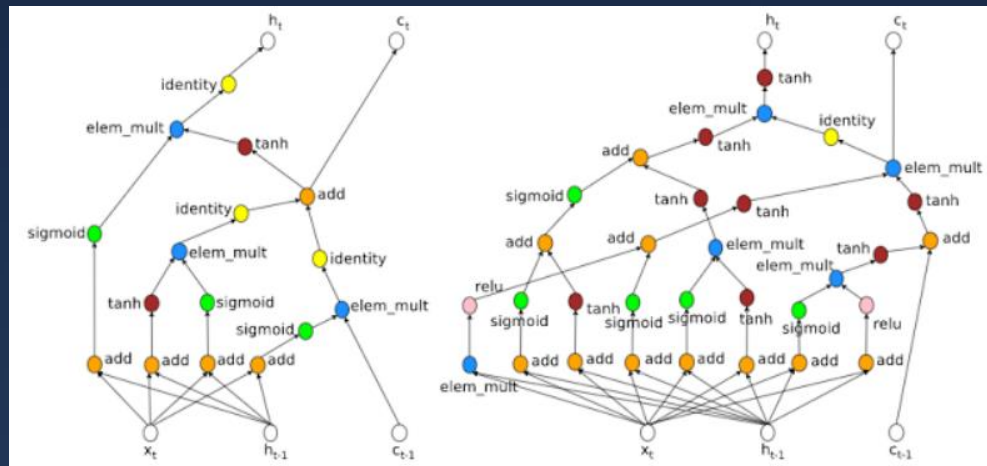
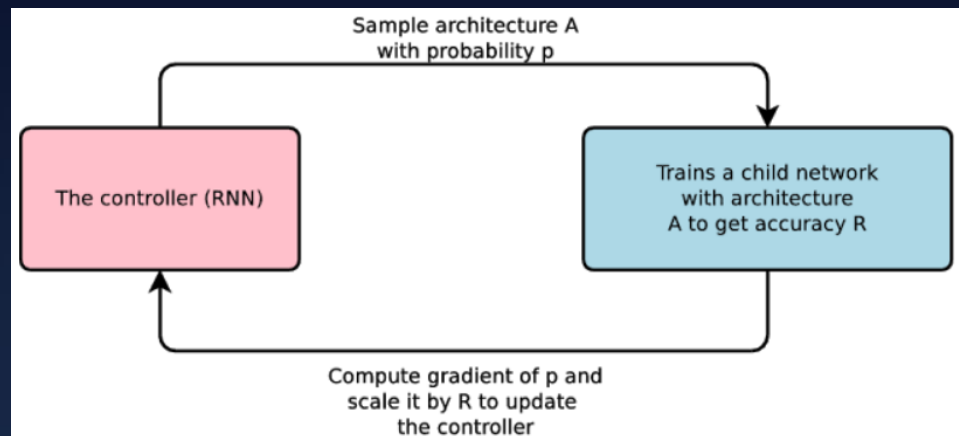
A job is started immediately upon submission. The running process is charged based on used resources.

Submit Job

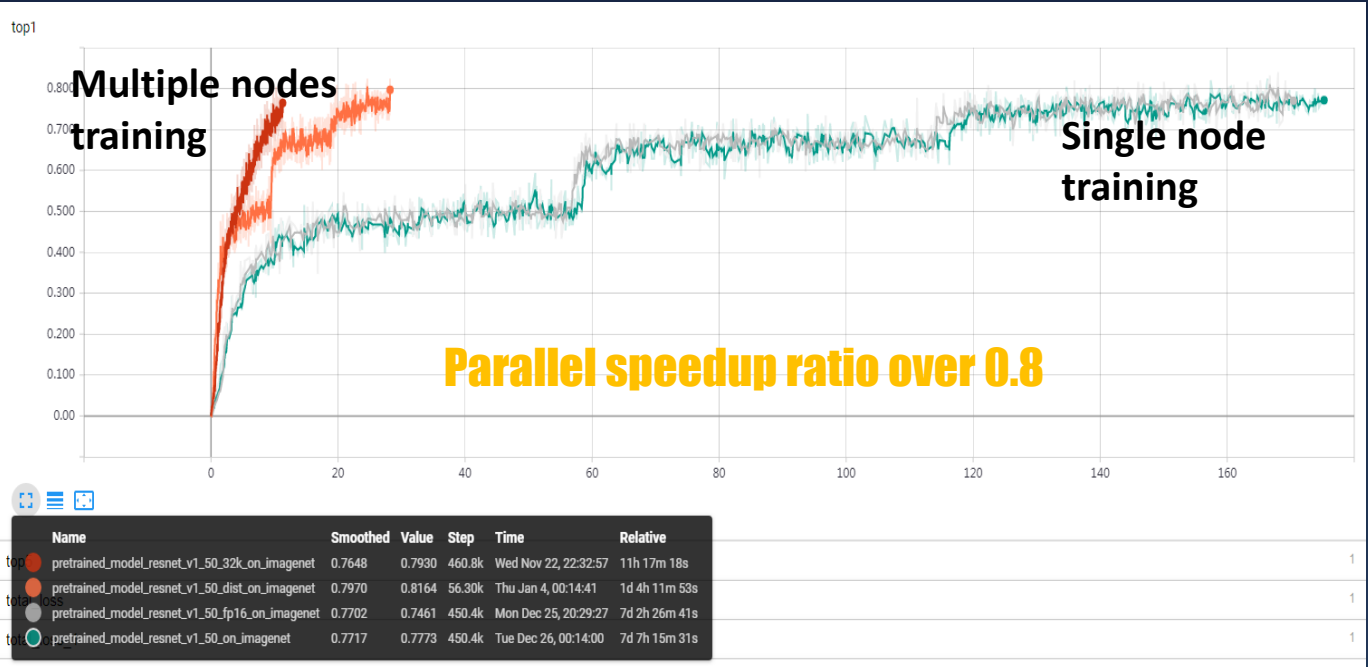
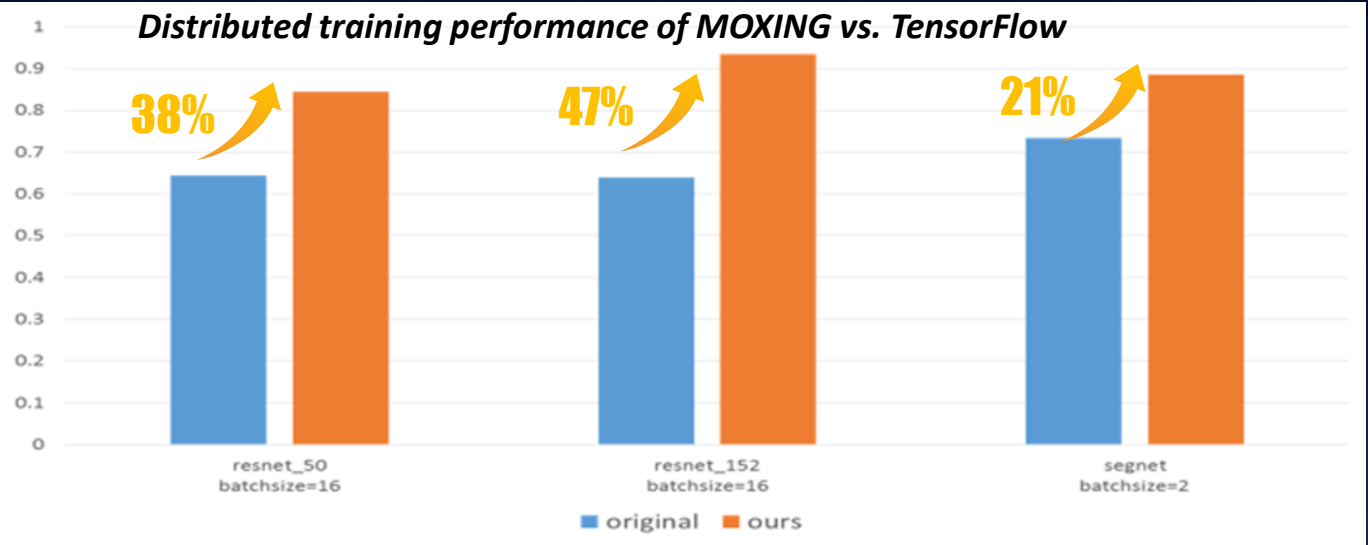
第1步：选择某个模型

第二步：配置某些参数（如路径等），Hyper Parameter可以自动被选择。

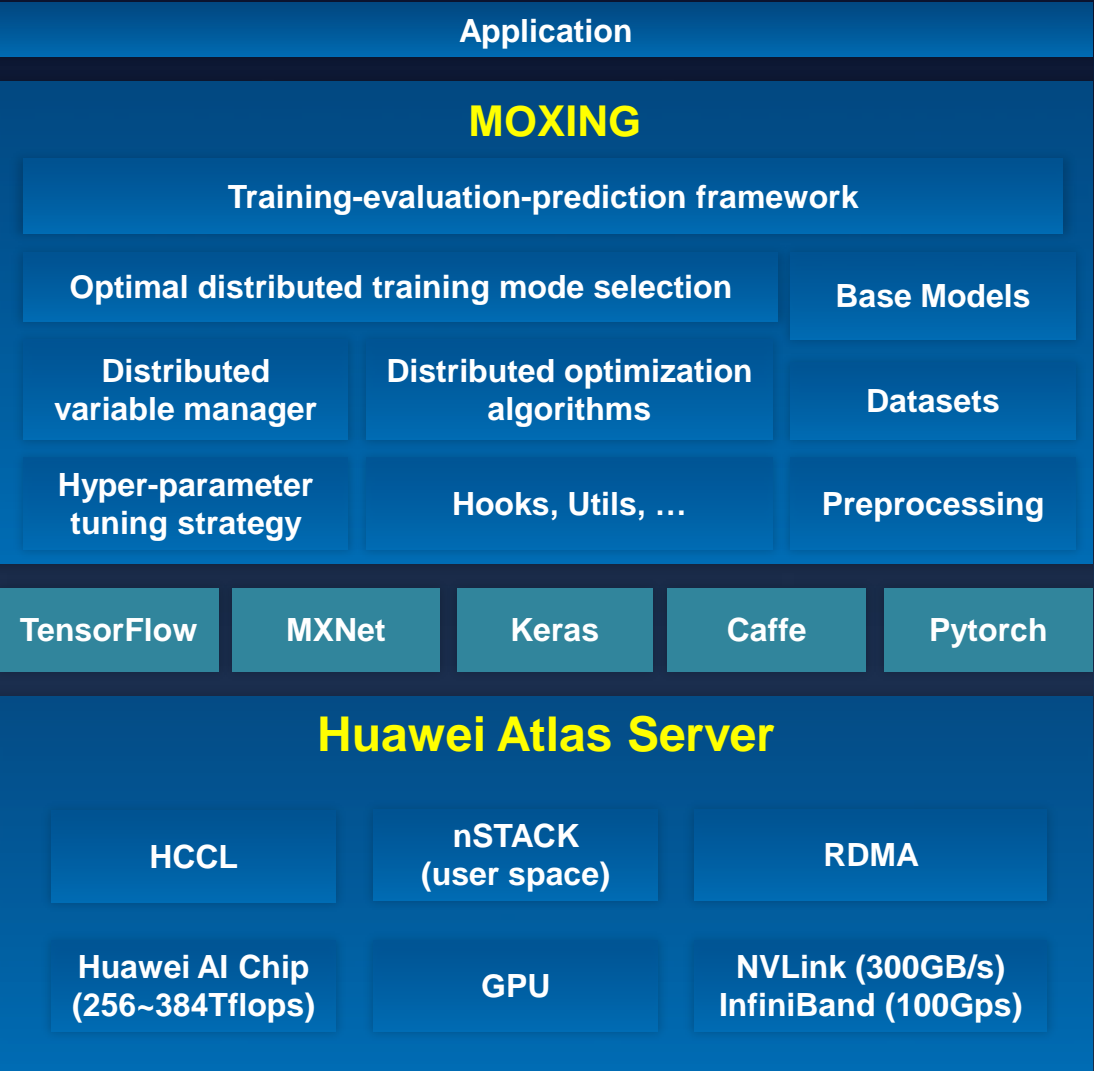
为应用开发者，提供自动模型设计



# 关键价值特性四：高性能、大规模分布式训练



**MOXING and Atlas Server (with Huawei AI Chip/GPU)**  
实现大规模、高性能分布式训练



# DLS服务使用页面

顶层菜单与功能



易用性与丰富功能的自然结合

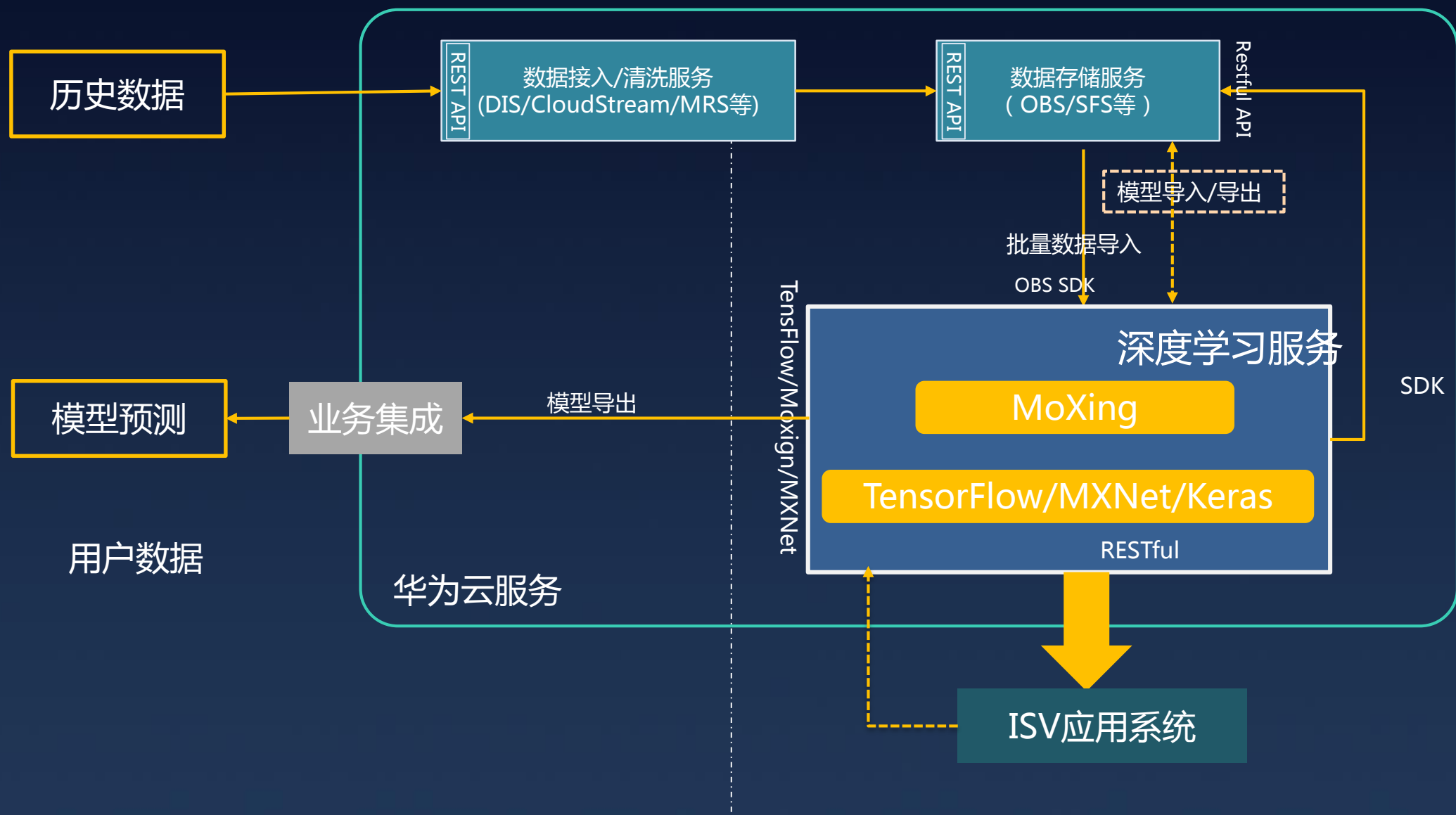
模型开发全流程、全周期平台支持

预制模型库支持快速实现模型训练

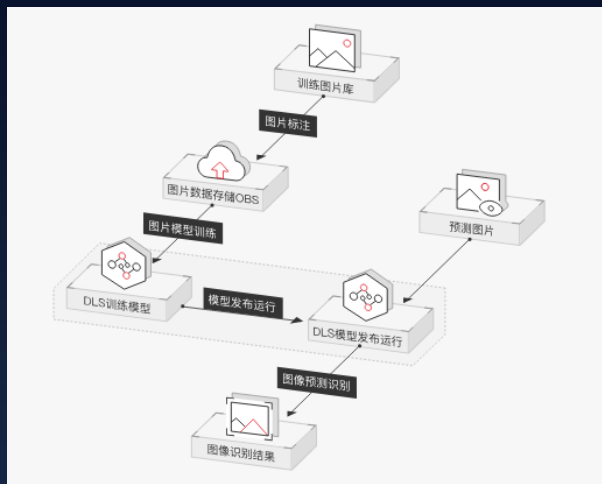
Notebook/TensorBoard等工具内置集成



# 深度学习服务服务整体解决方案

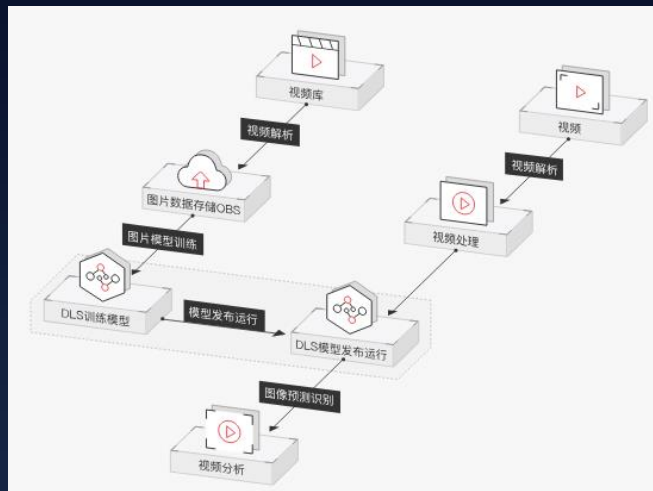


# 使用场景



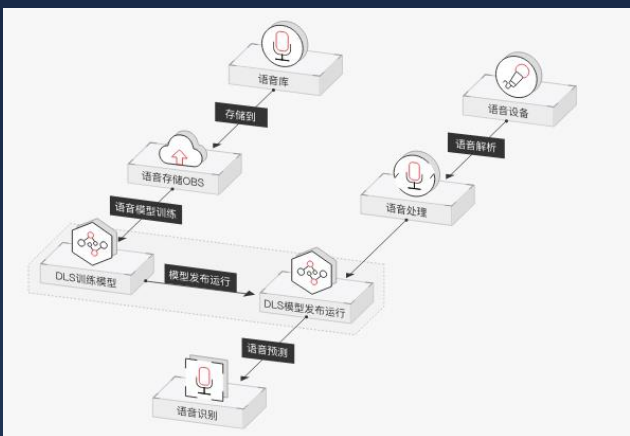
## 图像识别

准确识别图片中的物体分类信息，比如动物识别、品牌Logo识别、车型识别等



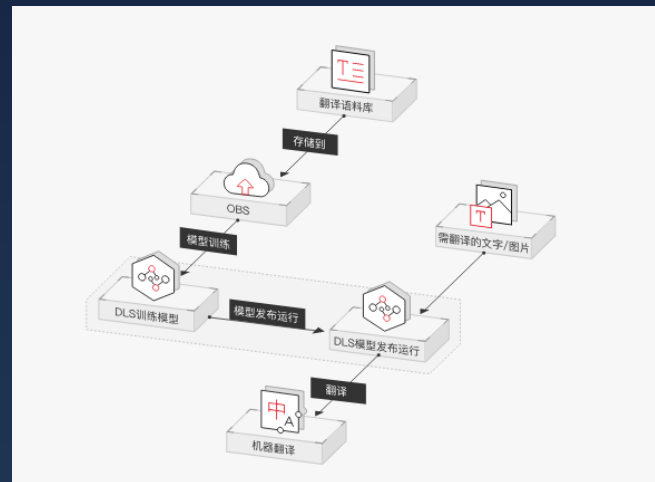
## 视频分析

准确分析到视频中的关键信息，比如人脸识别、车辆特征识别的场景应用



## 语音识别

可以让机器理解语音信号，协助处理语音信息，适用于智能客服问答，智能助手等



## 机器翻译

可以将翻译更加智能化，可以直接手机摄像头对准要翻译的文字，自动翻译成想要的语言文字





# Thank You.

**Copyright©2016 Huawei Technologies Co., Ltd. All Rights Reserved.**

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

## 华为云深度学习服务DLS