



萝卜快跑之案



端到端自动驾驶

关于自动驾驶内容

1. 端到端自动驾驶落地难点在哪?
2. 百度萝卜快跑与特斯拉的RobotTaxi 比较如何?



- 采访多个路人甲，对自动驾驶的看法（一本正经胡说八道）

01.

从萝卜快跑 到自动驾驶

萝卜快跑出圈时间线

1. 2022年8月，百度“萝卜快跑”在武汉和重庆正式开启车内无安全员的自动驾驶付费出行服务。
2. 长江日报2023年12月报道，截至2023年9月底，“萝卜快跑”已开通东西湖区、汉阳区等江北核心区域的贯通运营，在武汉接到逾34万个订单，是“萝卜快跑”全国城市之最。
3. 2023年12月13日，武汉市经信局的资料显示：武汉已成为全球最大的无人驾驶运营服务区，也是百度无人驾驶最大的运营服务区。
4. 2024年5月15日“Apollo Day”百度进一步迭代自动驾驶大模型，发布全球首个面向全自动驾驶的端到端大模型Apollo ADFM (Autonomous Driving Foundation Model)，L4自动驾驶。

新一代嘴王

- 2024年5月15日“Apollo Day”百度进一步迭代自动驾驶大模型，发布全球首个面向全无人自动驾驶的端到端大模型Apollo ADFM（Autonomous Driving Foundation Model），L4自动驾驶。王云鹏表示百度对“RobotTaxi”特别期待，也非常欢迎，不过也在安全方面，隔空向特斯拉下了战书：
- 特斯拉想要做到当前百度的全无人运营水平，还是需要挺长时间，少则3年、多则5载，如果不信可以先来武汉跑一跑。

新一代嘴王

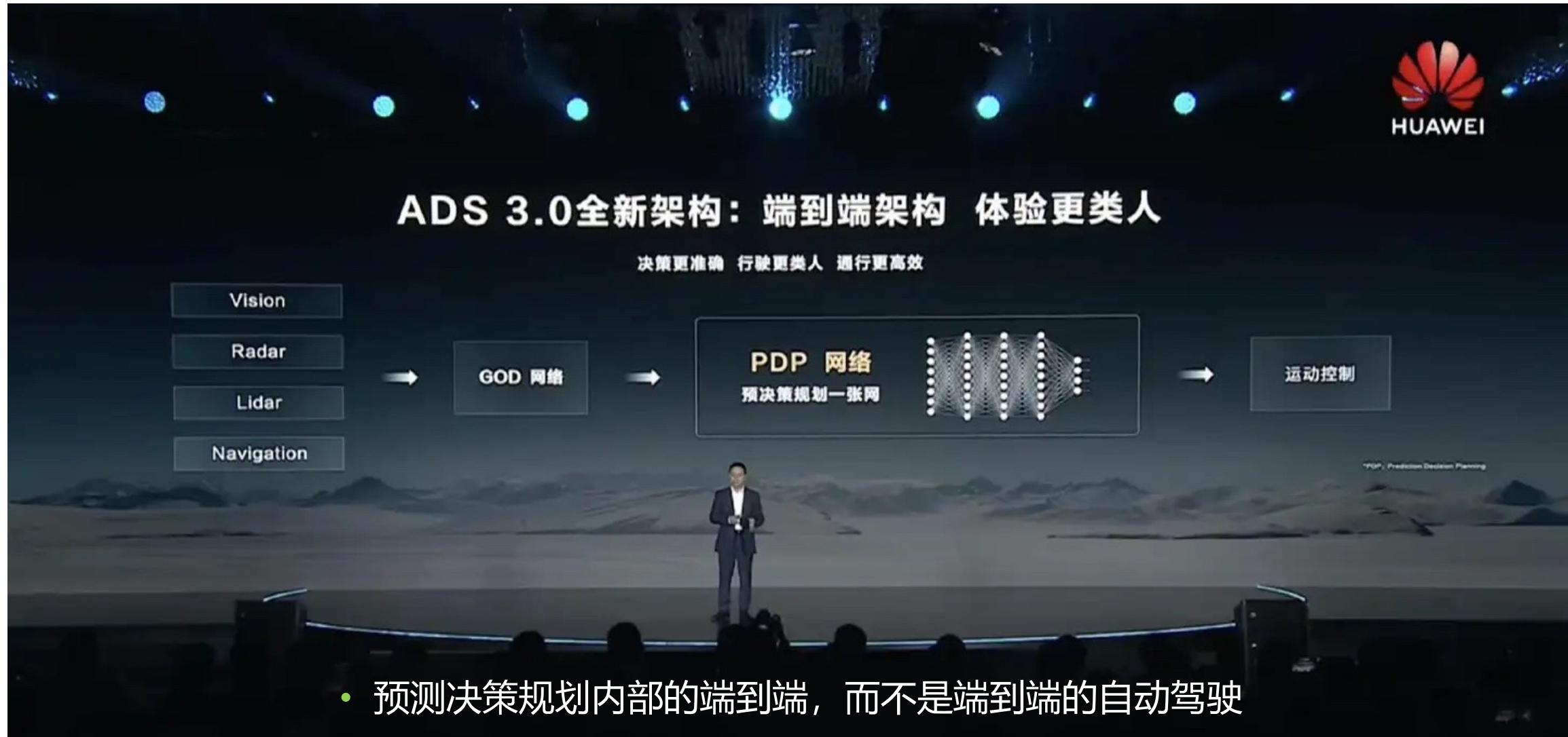
百度Apollo自动驾驶 比人类驾驶更安全

更安全的技术
1亿公里 0重大伤亡事故

更安全的产品
10重安全冗余 系统失效率低于 10^{-8}

更安全的运营
500万保险 出险率仅为人类司机的1/14

怎么可以少得了华为嘴哥



嘴王背后男人，始作俑者



**TESLA AUTOPILOT
FSD V12
DEMO DRIVE by ELON MUSK**



自动驾驶等级划分



- 通俗地，L1-L2级属辅助驾驶，目前国内大多数车企自动驾驶水平处于L2水平，司机不可或缺。
- L4绝大部分场景下，车辆可实现自动驾驶，不需要人类司机干预，即到L4级别，才可以称为自动驾驶。
- L5是在任何场景下，都可以实现自动驾驶。

搞清楚概念哦

- **萝卜快跑：**萝卜快跑自动驾驶出行服务平台（Robot 萝卜谐音梗）
- **百度阿波罗：**百度自动驾驶平台 Apollo， 提供智能驾驶解决方案及汽车智能化产品
- **特斯拉 FSD：**特斯拉 FSD (Full-Self Driving) 全自动驾驶解决方案

搞清楚概念哦

- **端到端自动驾驶**: 端到端自动驾驶
 - **大模型自动驾驶**: 用大模型来实现自动驾驶
 - **纯视觉自动驾驶**: 适用纯视觉作为输入的自动驾驶
-
- 端到端自动驾驶、大模型自动驾驶和纯视觉自动驾驶没有任何必然联系。三者之间有一些关联，但不等同。
- 三个概念完全独立存在，端到端自动驾驶系统不一定是传统意义上适用大模型驱动，也不一定就是纯视觉。

02.

自动驾驶 硬件平台

Apollo RT6：萝卜快跑使用的方案

- Apollo RT6：8个激光雷达、12个摄像头、6个毫米波雷达和12个超声波传感器，计算平台算力达到1200TOPS。
- 为了确保能够启动真正意义上无人驾驶运营，车型拥有电源、通讯、制动、转向、架构、计算单元和传感器七重冗余。



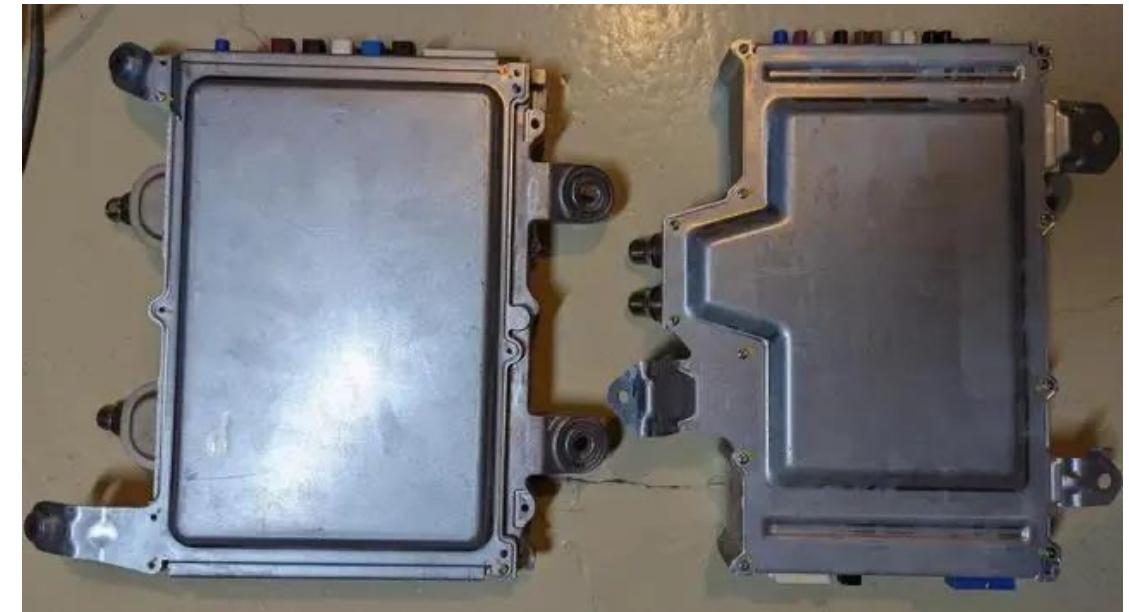
Apollo RT6：萝卜快跑使用的方案

- **感知系统：**配备了多颗高精度激光雷达、摄像头、毫米波雷达等传感器，形成了一个全方位的感知网络。这些传感器能够实时感知周围环境的动态变化，为车辆提供精确的位置和障碍物信息。
- **控制系统：**车辆内部的中央处理器（CPU）和高性能计算平台是无人驾驶汽车的大脑。它们负责接收传感器数据，进行实时处理和分析，并根据预设的算法生成相应的控制指令。
- **执行机构：**执行机构包括转向系统、制动系统、动力系统等，它们根据中央处理器的指令，精确地控制车辆的行驶状态。

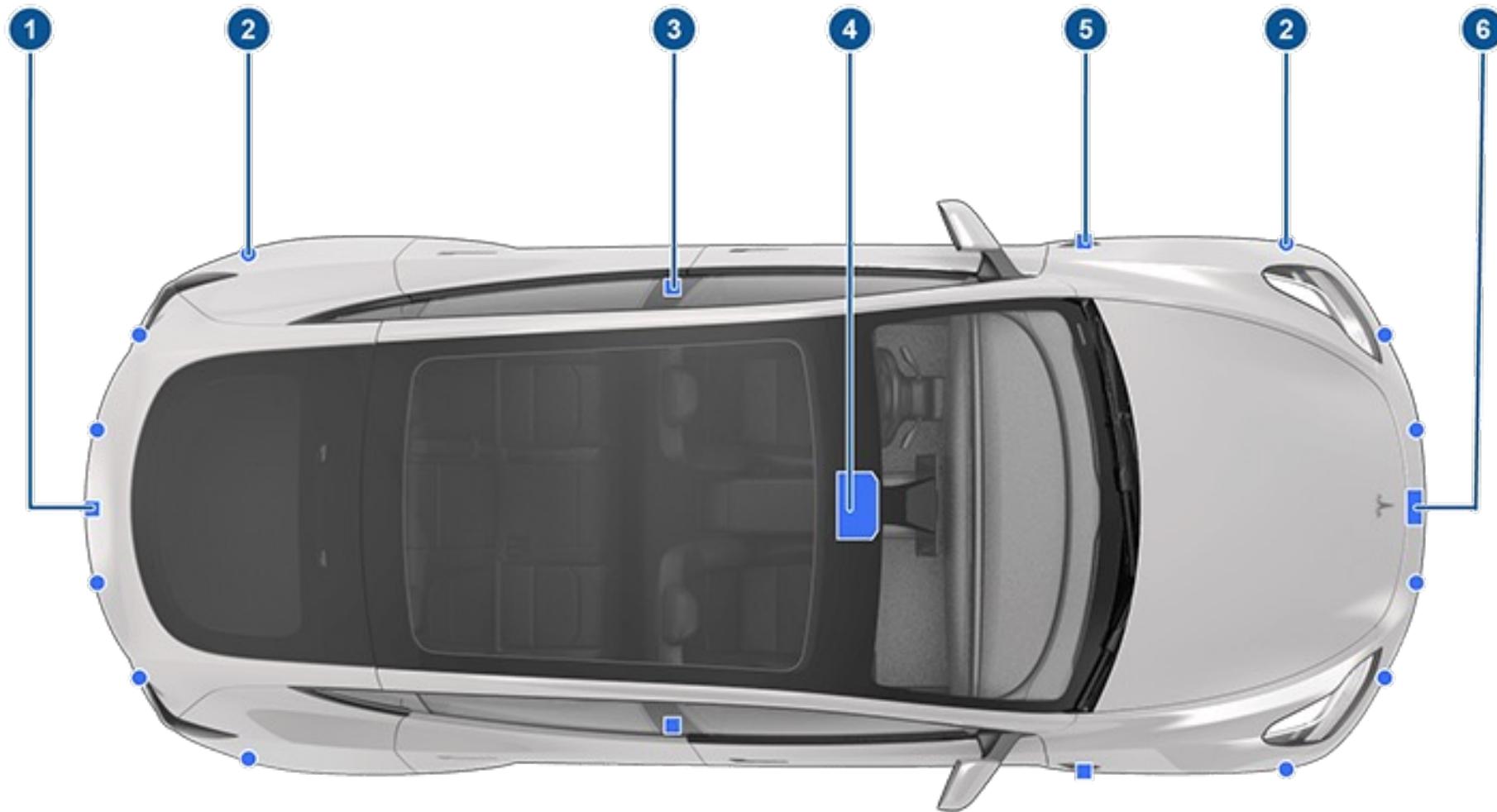


特斯拉 Model3 硬件

- 相比HW3.0，HW4.0略大并开始变得规整扁平化，安装点由原来的三个变成了四个，水冷的进出水口也有所调整。两代硬件接口也有很大变化，HW3.0原本有9个摄像头接口，而HW4.0增加到了12个，每个插口上都标记了摄像头的连接位置，比如F-SVC, L-SVC, R-FF-Rear等，其中一个标注为备用。



特斯拉 Model3 硬件



特斯拉 Model3 硬件

- "Simple design", no computing inside
 - Same CMOS Image Sensors



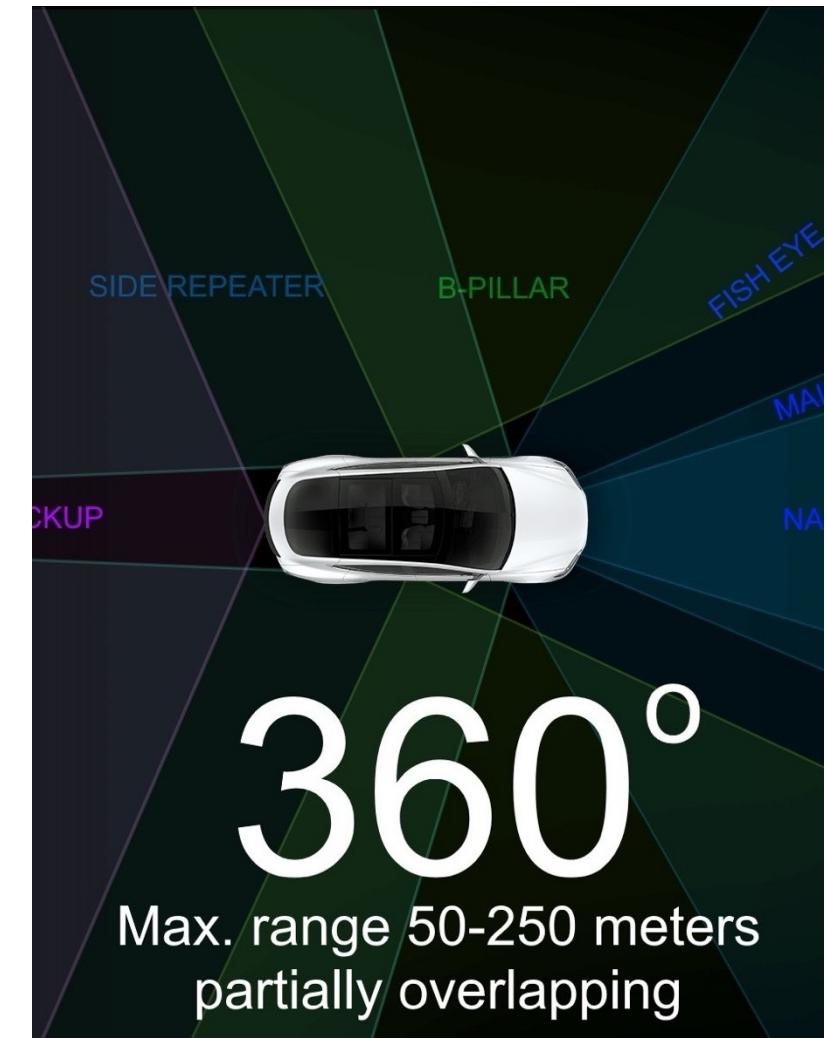
Courtesy of Tesla, 2020



Narrow (35°): 250 m
(820 ft) Main (50°): 150 m
(490 ft) Wide (120°): 60 m
(195 ft)



特斯拉 Model3 硬件



03. 级联 vs 端到端

传统自动驾驶系统

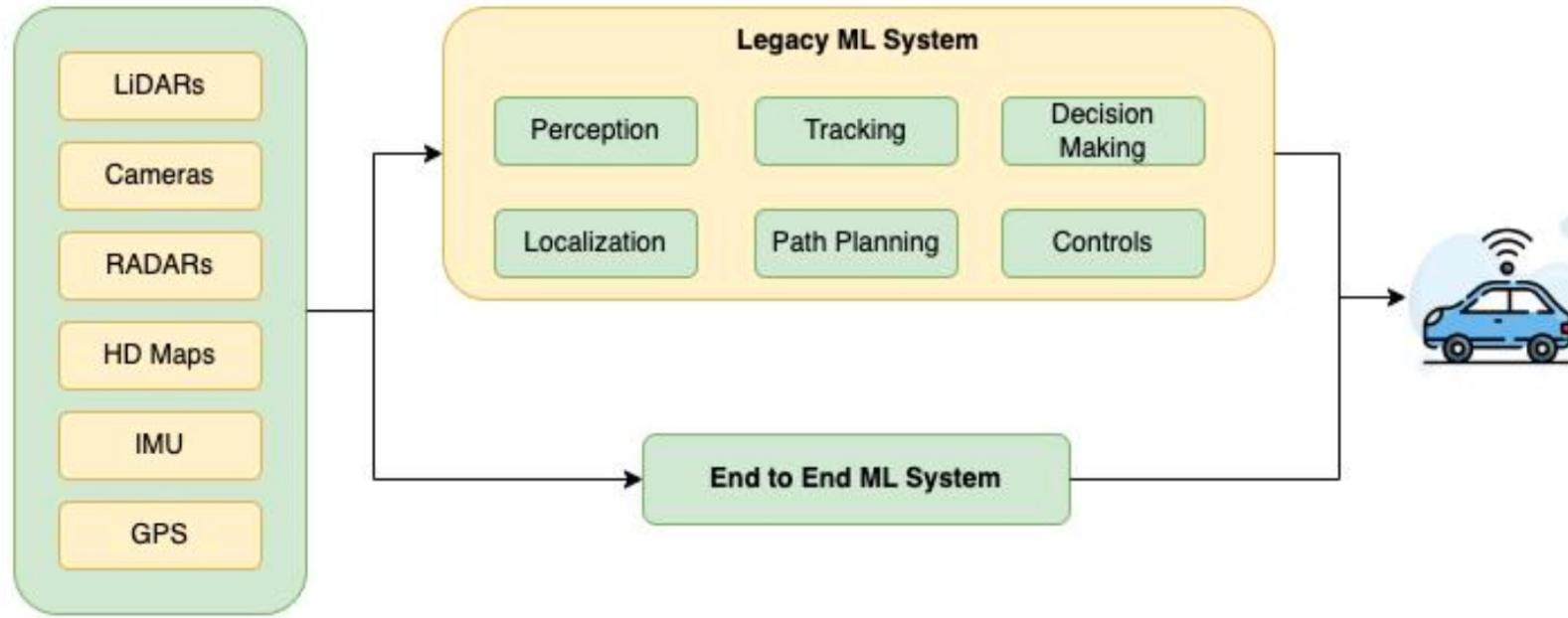
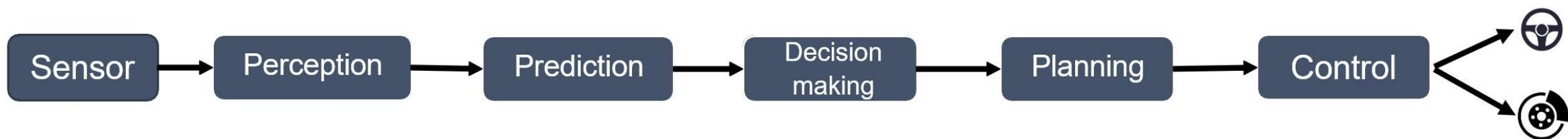


Fig. 2. Comparison of Modular Autonomous driving tasks vs single-model of end-to-end autonomous driving task.

1. 高精地图 (Mapping)
2. 感知 (Perception)
3. 定位 (Localization)
4. 目标跟踪 (Tracking)
5. 行为规划 (Predicting)
6. 路径规划 (Path Planning)
7. 决策 (Decision Making)
8. 控制 (Controls)
9. 仿真和测试 (Simulation)

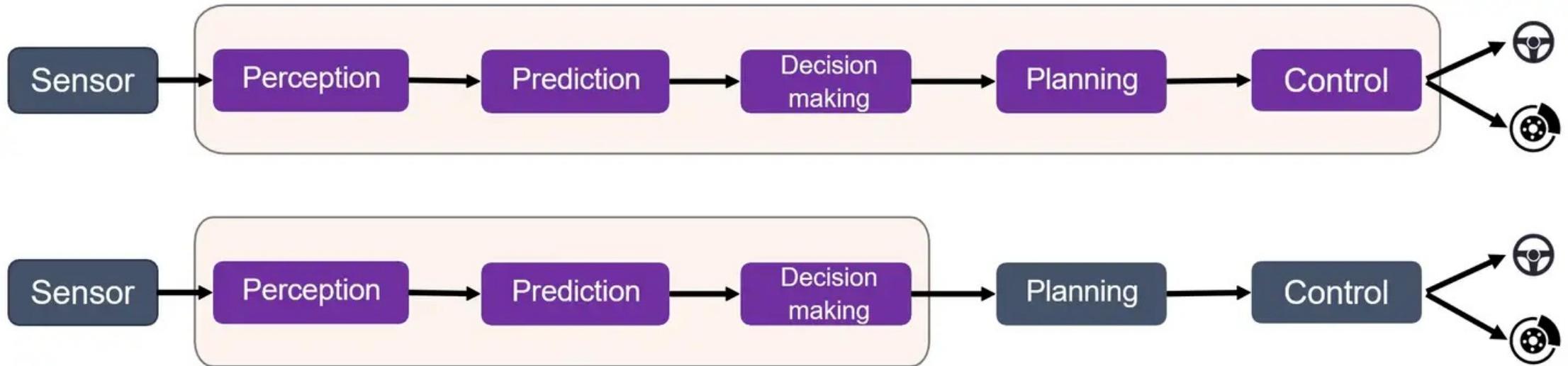
传统自动驾驶系统：级联式

- 每个模型都要专门进行训练、优化、迭代，随着模型不断进化，参数量提高和数据量增加，所需研发投入大，研发成本高。
- 模块化架构可以看做级联流水线，模型的输入参数，是前级模型的输出结果。如果前级模型输出结果有误差，会影响下一级模型输出，导致累计级联误差，最终影响系统性能。



端到端方式

- 通过一个模型实现流程中多个模型的功能。该模型可以接收传感器数据（图像、激光雷达等）作为输入，并输出车辆控制指令（如方向盘角度、刹车和加速等）。通过大规模数据集和训练算法，模型能够学习从感知到控制的完整驾驶策略。



级联方案与端到端方案对比

	级联	端到端
算法类型	模型算法 + 规则判断	模型算法 + 数据驱动
安全性	高	未知
可解释性	高	低
响应时延	中	高
算法难度	中	高
训练难度	中	高
评测手段	相同	相同
累计误差	高	无

04. 百度 vs 特斯拉

萝卜快跑 vs RobotTaxi

百度 Apollo ADFM

- Apollo ADFM 多模感知大模型，分为**感知大模型和规划决策大模型**两个部分。
 - 感知大模型包括检测、跟踪、理解、建图基本能力；
 - 规划决策大模型实现合规避障、博弈、预判等基本能力；
 - 链路模型综合输出多元环境信息，生成执行轨迹。
- 根据百度说法，基于大模型技术重构自动驾驶，可以兼顾技术的安全性和泛化性，做到安全性高于人类驾驶员10倍以上，实现城市级全域复杂场景覆盖。

百度 Apollo ADFM

- Apollo ADFM 多模感知大模型，分为**感知大模型**和**规划决策大模型**两个部分。
 - 感知大模型包括检测、跟踪、理解、建图基本能力；
 - 规划决策大模型实现合规避障、博弈、预判等基本能力；
 - **链路模型**综合输出多元环境信息，生成执行轨迹。
- 根据百度说法，**基于大模型技术重构自动驾驶**，可以兼顾技术的安全性和泛化性，做到安全性高于人类驾驶员10倍以上，实现城市级全域复杂场景覆盖。

百度 Apollo ADFM

第一代



第二代



第三代



第四代



第五代



第六代



2013年
BMW 3

2015年
奇瑞 EQ/比亚迪 秦/
北汽 EU260

2017年
林肯MKZ

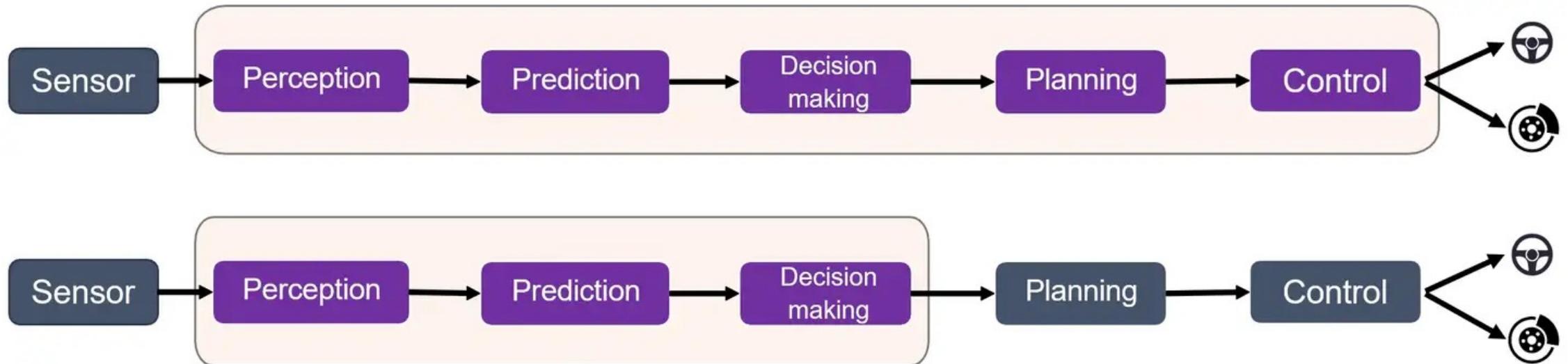
2019年
红旗E-HS3

2021年
Apollo Moon

2022年
Apollo RT6

Tesla FSD

- 端到端感知：输入是图像和传感器的数据，直接输出速度、tracking 结果，避免 Rule-based 的测速测距
- 感知到 Planning 的端到端：输入是图像和传感器的数据，直接输出方向盘、油门刹车的控制信号





[NO AUDIO]



STATE
Autonomous

MPH
16



4x

cruise



百度 vs 特斯拉 & 萝卜快跑 vs RobotTaxi

	百度萝卜快跑	特斯拉 RobotTaxi
算法	融合后级联 + 规则判断	端到端大模型 + 规则垫底
技术创新性	中	高
国内大面积铺开	低	低
可解释性	高	低
响应时延	中	高
算法难度	中	高
算力资源	中	高
累计误差	高	无

05. 技术思考

谁在 All in，谁在观望

- 理想汽车董事长兼 CEO 李想发表关于自动驾驶技术路线新思考：
 - 端到端+VLM（视觉语言模型）+生成式的验证系统，会是未来整个物理世界机器人最重要技术架构和技术体系
- 小鹏汽车宣布上车的端到端大模型由三部分组成：
 - 分别是神经网络 XNet（感知和语义），规控大模型 XPlanner 和大语言模型 XBrain（大场景认知）
- 特斯拉提出完全依赖神经网络端到端大模型：
 - 真正意义的端到端 AI，代码从 30 万行缩减到 2K，解决 corner case。
- 蔚来单独设立大模型部，负责端到端模型研发：
 - 原感知部门和规控部门下模型部合并，取消按功能（感知、地图、数据、规控等）划分。

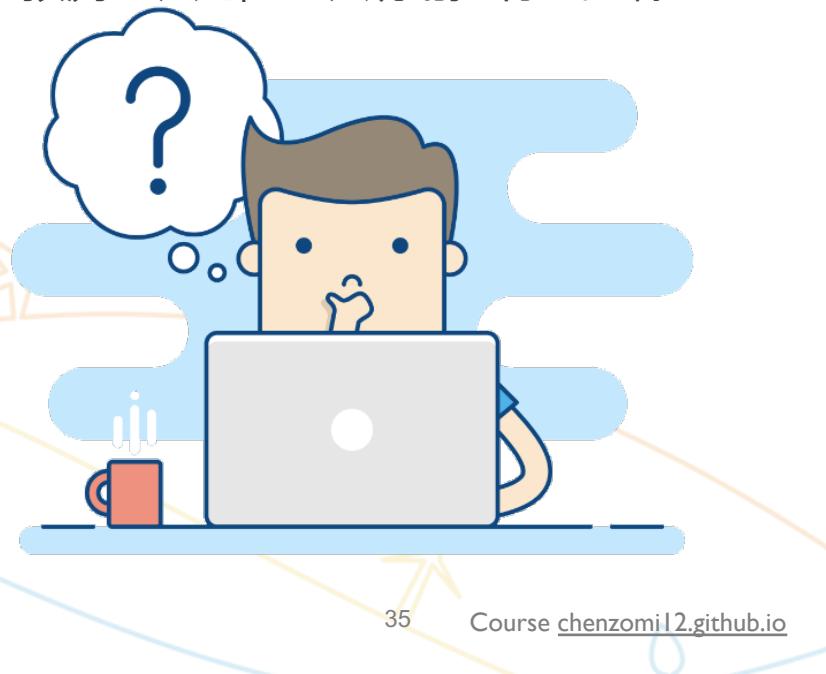
思考点

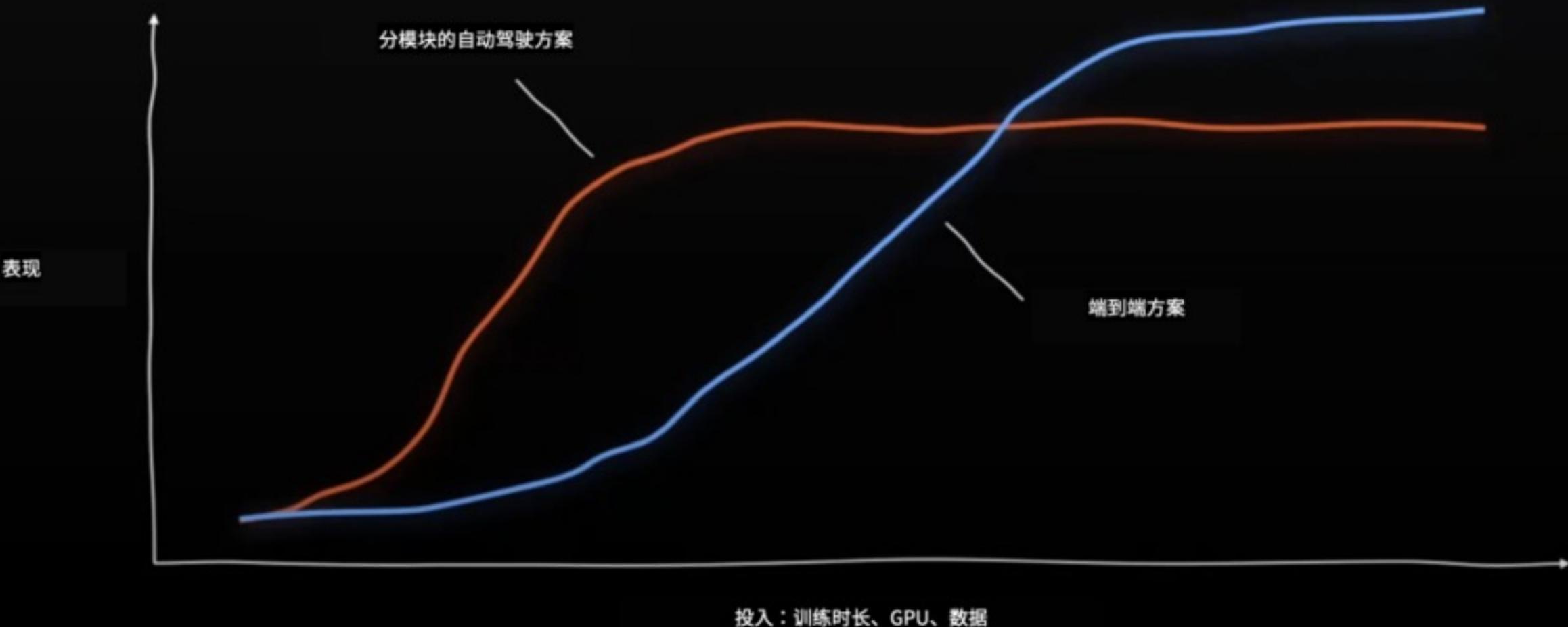
I. 端到端大模型最大难点是什么？

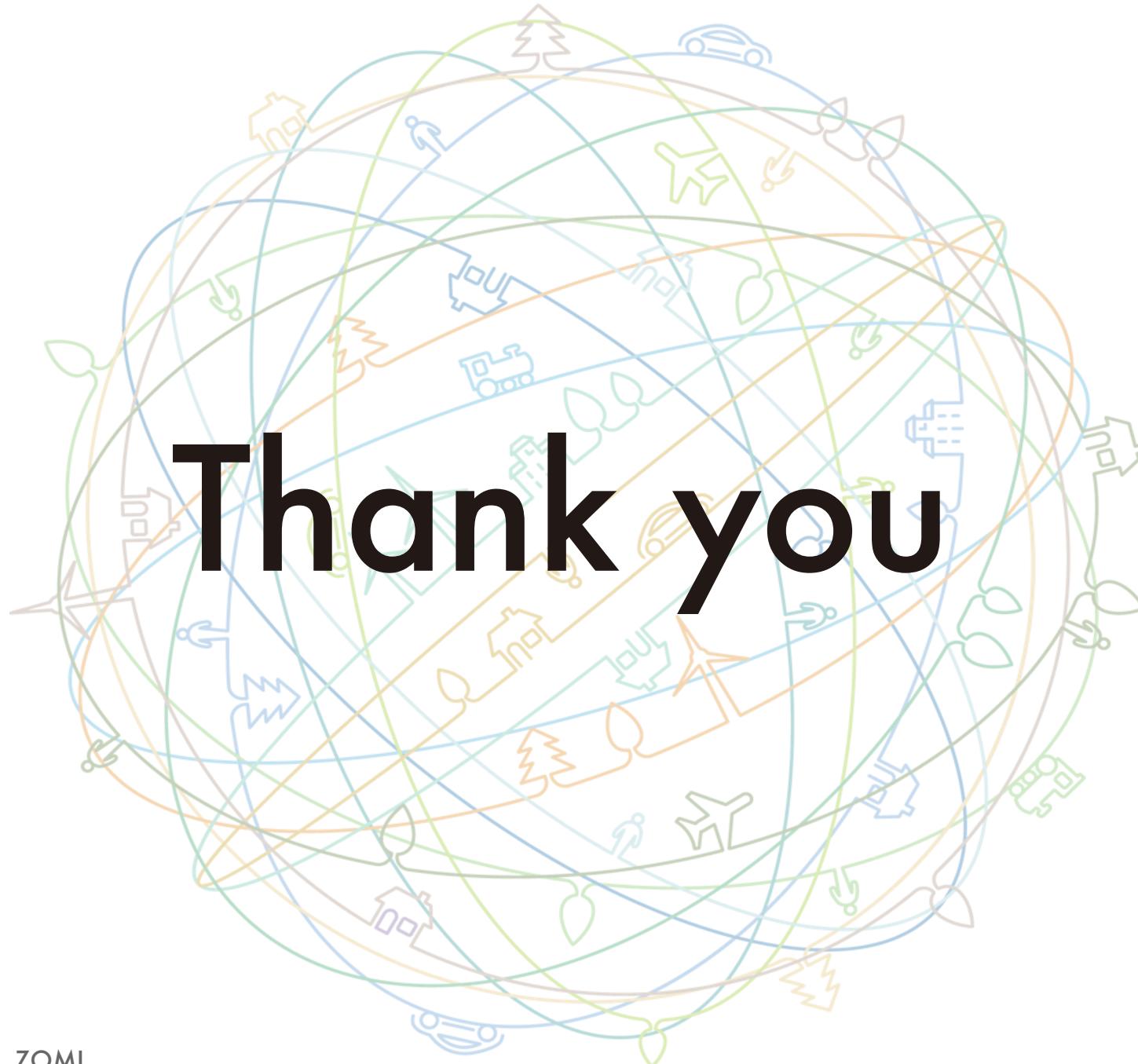
- 算力上，Musk表示目前 FSD 最大限制因素是算力，2024年Q1财报会议Tesla透露如今已有35000块H100计算资源，在2024年底将达到85000块。

2. 对组织如何变革？

- 目前大部分L4自动驾驶，自动驾驶团队组织架构以模块化为主，分为感知组、预测组、定位组、规划控制组。端到端技术架构直接打破不同模块间壁垒，需要整合全部来适应新技术范式。







Thank you

把AI系统带入每个开发者、每个家庭、
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2023 XXX 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. XXX may change the information at any time without notice.



Course chenzomi12.github.io

GitHub github.com/chenzomi12/DeepLearningSystem