

---

archi-intelligence Research Series

Working Paper 2026-03

---

# 多形态物理 AI 平台就绪度

Tesla FSD-Optimus Unified Stack

---

archi-intelligence Research Team

2026 年 6 月发布

[archi-intelligence.org](http://archi-intelligence.org)

---

## Front Matter

### F.3 版权与许可页

#### 版权信息 / Copyright Notice

本文档 © 2026 archi-intelligence Research Team。

本作品在 Creative Commons Attribution 4.0 International (CC-BY 4.0) 许可下发布。您可以自由：

- 分享——在任何媒介以任何形式复制和重新分发本作品
- 改编——修改、转换或基于本作品进行二次创作

唯一条件：必须以适当方式标明出处、提供许可链接、说明是否做出了修改。

License URL: <https://creativecommons.org/licenses/by/4.0/>

#### 数字对象唯一标识符 / DOI

DOI: 10.5281/zenodo.20513903 Permanent Archive: <https://doi.org/10.5281/zenodo.20513903>

#### 引用格式建议 / Suggested Citation

archi-intelligence Research Team. (2026). Multi-Embodiment Physical AI Platform Readiness: The Tesla FSD-Optimus Unified Stack (Working Paper 2026-03). archi-intelligence Research Series. <https://doi.org/10.5281/zenodo.20513903>

#### BibTeX 格式

```
@techreport{archi_intelligence_2026_03,  
  title = {Multi-Embodiment Physical AI Platform Readiness:  
          The Tesla FSD-Optimus Unified Stack},  
  author = {{archi-intelligence Research Team}},  
  institution = {archi-intelligence},  
  year = {2026},  
  number = {2026-03},  
  type = {Working Paper},  
  series = {archi-intelligence Research Series},  
  url = {https://archi-intelligence.org/research/2026-03},  
  doi = {10.5281/zenodo.20513903}  
}
```

#### 联系方式 / Contact

- 研究团队 / Research Team: [research@archi-intelligence.org](mailto:research@archi-intelligence.org)
- 媒体咨询 / Press Inquiries: [press@archi-intelligence.org](mailto:press@archi-intelligence.org)
- 内容更正 / Corrections: [corrections@archi-intelligence.org](mailto:corrections@archi-intelligence.org)
- 网站 / Website: <https://archi-intelligence.org>

### F.4 利益声明 (COI Disclosure)

#### 利益冲突与编辑独立性声明

为保持本研究系列的学术独立性与公信力，archi-intelligence Research Team 在此明确以下事项。

##### 1. 关于本研究机构

archi-intelligence 是致力于将 Architecture Intelligence (架构智能, AI<sup>2</sup>) 建立为研究范式的独立学术研究机构。本机构的使命是通过公开方法论、透明数据归因、严格同行评审，推动跨产业架构工程实践的标准化与可比性。

## 2. 关于资金与赞助

本研究系列在本次发布阶段，未接受任何被评估对象（包括但不限于本报告中提及的 OEM、Tier-1 供应商、平台公司、芯片厂商）的直接或间接资金支持。

archi-intelligence 在初始阶段由 Arkimind（一家专注于汽车电子电气架构智能化的商业实体）提供 seed sponsorship。这一关系在本声明中明确披露，并通过以下治理隔离机制保障编辑独立性：

- 编辑/研究团队不向 Arkimind 销售或业务团队汇报
- 研究人员的薪酬不与 Arkimind 任何商业指标挂钩
- 报告发布前，Arkimind 商业团队无审阅或修改权
- 评估方法论与原始数据来源 100% 公开
- 任何被评估对象可提交修正反馈，处理流程公开

## 3. 关于编辑独立性

本研究的方法论选择、案例评估、结论表述均由 archi-intelligence Research Team 独立做出。研究判断不受任何商业利益、政治立场、地缘政治偏好的影响。我们承认这一编辑独立性最终需要由读者通过持续审视我们的研究质量来验证。我们欢迎来自学术界、产业界、监管机构的批评性反馈，并承诺在后续版本中公开记录所有重要修正。

## 4. 关于数据来源与时间窗口

本研究采用 2024 年 1 月至 2026 年 5 月期间的公开来源信息，包括但不限于：上市公司财报与 SEC/HKEX 公开文件（含 Tesla Q1 2026 Update、SpaceX S-1 招股书 2026.5.20、Xpeng SEC Form 6-K、Xiaomi HKEX 年报等 Tier 1 资料）；OEM 与供应商官方技术发布（AI Day、HDC、Investor Day、CES、Tesla Annual Shareholder Meeting 2025.11.6 等）；已公开的专利文献；同行评审学术论文与会议出版物；行业标准组织发布（AUTOSAR、ISO、SAE、IEEE 等）；主流技术媒体报道与产业分析。对于无法独立验证的传闻信息或匿名来源信息，本研究采取保守立场——要么不引用，要么明确标注为“基于公开报道的方向性推断”。

对于马斯克的公开声明与社交媒体发言，本报告区分两类引用：（a）作为官方文件或财报电话会议正式记录的，按 Tier 2 处理；（b）作为社交媒体非正式发言的，按 Tier 4 处理并明确标注“非正式来源”。马斯克的许多性能声明（如“AI5 比 AI4 快 40 倍”）为综合性、营销性口径，本报告标注其来源性质，不作为客观比较基准采用。

## 5. 关于研究限制

作为单一案例的深度研究，本工作公开承认：（a）Tesla 案例处于高速演进期，本报告数据截止 2026 年 5 月 21 日，所述事实可能在数月内被新事件刷新；（b）Tesla 通过 AI Day、各季度财报、SEC 文件、SpaceX S-1 招股书提供了相对丰富的技术披露，但仍有大量底层实现细节（如 FSD 端到端网络精确架构、Cortex 2.0 具体拓扑、AI5 微架构）未公开；（c）单一案例研究不代表对 AR4 路径的完整覆盖——华为 HIMA、小鹏等其他 AR4 候选案例需独立研究。本报告在第九章对 5 家镜像 OEM 的分析，是对单案例局限的部分补偿，但不能替代对每个案例的独立深度研究。

## 6. 关于商标与产品名

本报告中提及的所有商标、产品名、产品代号（包括但不限于 Tesla、Optimus、FSD、SpaceX、Anthropic、Volkswagen、Toyota、Huawei、Xpeng、Xiaomi 等）均归各自权利人所有。引用此类名称纯粹用于学术讨论与产业分析，不构成任何商业关联、背书或推广。

## 7. 关于本报告的 AI 辅助披露

本报告的部分草稿生成与文本润色，使用了 Anthropic 公司的 Claude 模型作为辅助工具。最终的研究方法论选择、事实核实、判断推导、结论表述，均由 archi-intelligence Research Team 独立做出，并通过分级证据链 Tier 1-4 严格回溯。

需要透明披露的相关商业关系：根据 SpaceX 于 2026 年 5 月 20 日提交 SEC 的 S-1 招股书（本报告

Tier 1 资料)，Anthropic 与 SpaceX 之间存在云服务商业关系（Anthropic 向 SpaceX 支付 \$1.25B/月至 2029 年 5 月，使用 COLOSSUS / COLOSSUS II 部分容量）。SpaceX 与 Tesla 在 2026 年 3 月就 Terafab 项目达成框架协议（同样为 Tier 1 公开披露）。这构成”Anthropic ← SpaceX ← Musk → Tesla”的间接商业链条。

本报告主体涉及 Tesla 案例研究，因此 archi-intelligence 在此明确：（a）archi-intelligence 不接受 Anthropic、SpaceX、Tesla、xAI 中任何一方的直接或间接资金支持；（b）使用 Claude 作为草稿工具的事实，不影响本报告的方法论独立性与编辑判断；（c）本报告对 Tesla 的批判性审视（包括 HW3 撤回事件、AI5 不给车的资源配置、Robotaxi 部署落差、CEO 个人风险、scaling laws 持续性等）已充分体现编辑独立性。

## 8. 免责声明

本报告不构成任何形式的投资建议、商业咨询或法律意见。任何基于本报告内容所做的商业决策，由决策方自行承担风险与责任。

- 研究联系：research@archi-intelligence.org
- 内容更正与反馈：corrections@archi-intelligence.org
- 完整治理章程：<https://archi-intelligence.org/governance>

## Contents

Front Matter	1
F.3 版权与许可页	1
F.4 利益声明 (COI Disclosure)	1
F.5 详细目录	4
F.6 图表索引	6
F.7 缩略语表	7
摘要	10
引言: 为什么 Tesla 是 AR4 垂直闭环的唯一完整案例	11
本研究在 archi-intelligence Research Series 中的位置	11
Tesla 选题的方法论判断	11
本报告的研究边界	11
本报告的结构	12
第一章 硅路线: 从 HW3 到 AI7 的代际演进	13
1.1 代际总览	13
1.2 HW3: 自研推理芯片的起点 (2019) 与无监督承诺的撤回	14
1.3 HW4 / AI4: 当前量产主力 (2023) 与 AI4.5 的 stopgap 浮现	14
1.4 AI5 / HW5: 跨代际跃升 (2026.4 流片) 与”不给车”的硬件分级	15
1.5 AI6 / AI7: 未来路标与 Terafab 计划	15
1.6 硅路线的架构启示	16
第二章 FSD 软件栈演进: 从规则到端到端的范式革命	17
2.1 V11 及之前: 模块化规则栈的黄昏	17
2.2 V12: 端到端革命 (2024.1)	18
2.3 V13: HW4 原生分辨率 (2024.12)	18
2.4 V14: 多模态与世界模型 (2025-2026)	18
2.5 Robotaxi 部署与监管现实	18
2.6 软件栈演进的架构启示	19
第三章 占用网络与形态无关感知	20
3.1 占用网络的技术本质	20
3.2 为何占用网络是”形态无关”的	20
3.3 占用网络在 Tesla 感知栈中的位置	21
3.4 占用网络的限制与争议	21
3.5 占用网络的架构启示	21
第四章 跨形态复用机理: 从汽车到人形机器人	22
4.1 跨形态复用的官方披露: Tesla 11 层共享核心技术	22
4.2 11 层的三组结构: D3 提炼的三组复用机理	23
4.3 智能层 5 项的工程机理	24
4.4 物理层 6 项的累积逻辑	24
4.5 组织层: 隐性的第 12 层	25
4.6 跨形态复用的真实成本结构与不可复制性	25

第五章 Optimus 运动学与不可复用层：差异的工程颗粒度 . . . . .	28
5.1 “大脑”可复用，“小脑”不可复用 . . . . .	28
5.2 Optimus 的运动学规格 . . . . .	28
5.3 不可复用层的工程内容 . . . . .	28
5.4 可复用与不可复用的边界 . . . . .	28
5.5 不可复用层的战略含义 . . . . .	29
第六章 训练基础设施：Dojo → Cortex → Dojo 3 的演化 . . . . .	31
6.1 Dojo 的兴衰（2019-2025） . . . . .	31
6.2 Cortex：务实的混合方案（2024-） . . . . .	31
6.3 Cortex 2.0 与 Dojo 3 重启（2026） . . . . .	31
6.4 训练基础设施演化的架构启示 . . . . .	32
6.5 SpaceX-xAI 合并后的集团级算力重构（2026.2-2026.5） . . . . .	32
第七章 组织层面的复用：FSD 与 Optimus 团队合并的架构含义 . . . . .	33
7.1 团队合并的事实 . . . . .	33
7.2 康威定律的逆向应用 . . . . .	33
7.3 对”既造车又造机器人”公司的启示 . . . . .	33
7.4 组织复用的代价 . . . . .	33
第八章 AR4 的代价与边界：对其他玩家的启示 . . . . .	34
8.1 AR4 垂直闭环的五重代价 . . . . .	34
8.2 对其他玩家的三类启示 . . . . .	34
8.3 跨形态复用的边界：诚实的总结 . . . . .	35
8.4 Tesla 作为 AR4 参照系的最终意义 . . . . .	35
第九章 镜像章：D3 对 D2 其他 OEM 的论证意义 . . . . .	37
9.1 大众 CARIAD：全球总部 vs 中国速度的双线张力 . . . . .	37
9.2 丰田 Arene：渐进式 + fail-operational 与 Tesla 的范式对立 . . . . .	38
9.3 华为 HIMA：AR4 第二路径（ICT 进入汽车 vs 车企走向 ICT） . . . . .	38
9.4 小鹏：最像 Tesla 的中国新势力，晚 5-10 年 . . . . .	39
9.5 小米：消费电子供应链效率 + 跨界进入的结构短板 . . . . .	41
9.6 5 家镜像综合：基于 12 层的可复用矩阵 . . . . .	42
第十章 实证章：D3 对 D1 核心观点的支撑与挑战 . . . . .	44
10.1 两层结构论的验证：基础设施收敛 vs 控制语义发散 . . . . .	44
10.2 “AI 是语言能力，非大脑”在 Tesla 端到端的检验 . . . . .	44
10.3 架构债务理论的 Tesla 实证 . . . . .	45
10.4 失效哲学：fail-soft vs fail-operational 的 Tesla 定位 . . . . .	45
10.5 工具范畴的反向定义：追赶者需要什么样的方法论与工具 . . . . .	46
10.6 D1 的可证伪点：Tesla 路径未来可能的失败模式 . . . . .	46
结论：AR4 作为参照系的方法论意义 . . . . .	48
Tesla 作为参照系的三重价值 . . . . .	48
AR4 参照系的边界 . . . . .	48
致追赶者 . . . . .	49
致 D1 的可证伪承诺 . . . . .	49
附录 A: Tesla AR4 关键事实时间线(Tier 1-2 来源) . . . . .	50
附录 B: 关于 archi-intelligence 研究系列 . . . . .	51

## F.6 图表索引

### 图 (Figures)

- 图 1.1 Tesla 硅代际演进时间线 (HW3 → AI4 → AI4.5 → AI5 → AI6 → AI7 + Terafab)
- 图 2.1 FSD 软件栈范式演化 (V11 → V14 + 软件能力 ≠ 部署能力)
- 图 4.1 跨形态复用 12 层栈 (Tesla 11 层共享 + 组织层第 12 层)

### 表 (Tables)

- 表 1.1 Tesla 自动驾驶硅代际对照 (截至 2026 年 5 月)
- 表 4.1 Tesla 官方披露的 11 层共享核心技术 (2025.11.6 股东大会)
- 表 4.2 D3 对 Tesla 11 层共享技术的三组结构提炼
- 表 9.1 5 家追赶者 × 12 层可复用矩阵

## F.7 缩略语表

### 架构等级与框架（与 D1/D2 系列一致）

- **AR0-AR5** — Architecture Readiness 0-5, 架构成熟度阶梯（首篇旗舰报告提出）
- **AR4** — Multi-Embodiment Physical AI Platform（多体物理 AI 平台；本报告核心研究对象）
- **AI<sup>2</sup>** — Architecture Intelligence（架构智能；本研究系列的研究范式）
- **AI<sup>2</sup>-ML** — Architecture Intelligence Maturity Levels（架构智能成熟度等级；L0-L5）

#### Tesla 硅 / 软件代际

- **HW3** — FSD Computer（2019 年量产, 14nm Samsung, ~144 TOPS）
- **HW4 / AI4** — 当前量产主力（2023 年起, Samsung 7nm, ~500 TOPS）
- **AI4.5** — stopgap chip（因 AI5 延期, 2025 年末搭载于 2026 Model Y）
- **HW5 / AI5** — 2026.4 流片, ~4000 TOPS 等效（5× AI4 useful compute）
- **AI6 / AI7** — 未来代际路标
- **FSD** — Full Self-Driving（Tesla 自动驾驶软件包）
- **V11 / V12 / V13 / V14** — FSD 软件栈代际
- **VLA** — Vision-Language-Action（视觉-语言-动作模型）

#### Tesla 集团级算力基础设施

- **Cortex 2.0** — Tesla 训练集群（Giga Texas, ~100K H100/H200, 250-500MW）
- **COLOSSUS / COLOSSUS II** — xAI / SpaceX 训练集群（Memphis TN + Southaven MS, ~1 GW）
- **Dojo** — Tesla 自研训练芯片（D1 / Dojo 3）
- **Terafab** — SpaceX + Tesla + Intel 联合芯片制造项目（\$25B Austin, 长期目标 1 TW/年）
- **xAI Merger** — SpaceX 收购 xAI（2026.2.2, 合并估值 \$1.25T）

#### Tesla / Optimus 物理硬件

- **Cybercab** — Tesla 专用 Robotaxi 车型（Q2 2026 SOP, 初期使用 AI4/AI4.5）
- **Optimus** — Tesla 人形机器人
- **Optimus Gen 3** — 手部升级版（22 DoF/手 + 50 执行器, 腱驱动）
- **4680** — Tesla 自研锂电池电芯规格
- **Gigacasting** — Tesla 巨型铸件制造工艺

#### FSD 软件栈关键概念

- **End-to-End** — 端到端神经网络（V12 起, “Photon In, Control Out”）
- **Occupancy Network** — 占用网络（3D 体素表示, 形态无关感知）
- **Software 2.0** — 数据驱动而非规则手写的软件范式
- **Robotaxi** — Tesla unsupervised FSD 商业运营服务
- **Intervention-free Streak Counter** — V14.3 引入的”无干预连续里程”计数器

#### 安全与监管

- **ASIL** — Automotive Safety Integrity Level（汽车安全完整性等级, ISO 26262）
- **SOTIF** — Safety of the Intended Functionality（预期功能安全, ISO 21448）
- **UN-R155** — 网络安全管理体系（CSMS）联合国法规
- **UN-R156** — 软件更新管理体系（SUMS）联合国法规
- **UN-R157** — 自动车道保持系统（ALKS）联合国法规
- **L2 / L3 / L4 / L5** — SAE J3016 自动驾驶等级
- **fail-soft** — 软失效（系统失效后由人类作为安全冗余）
- **fail-operational** — 容错运行（任何子系统失效后系统自身可降级安全完成任务）
- **NHTSA** — National Highway Traffic Safety Administration（美国国家公路交通安全管理局）

- **RDW** — Rijksdienst voor het Wegverkeer (荷兰车辆管理局)
- **SEC 文件与官方资料**
- **10-Q** — 美国上市公司季度财报 (SEC Form)
- **8-K** — 美国上市公司重大事件公告 (SEC Form)
- **S-1** — 美国 IPO 招股书 (SEC Form)
- **6-K** — 外国发行人定期或临时报告 (SEC Form)
- **HKEX** — Hong Kong Stock Exchange (香港交易所)
- **镜像 OEM 缩略**
- **CARIAD** — 大众集团软件子公司
- **CARIZON** — 大众与地平线合资公司 (60:40)
- **CEA** — China Electronic Architecture (中国电子电气架构; VW 大众)
- **SSP** — Scalable Systems Platform (大众全球可扩展系统平台)
- **VCTC** — Volkswagen Group China Technology Company
- **Arene** — Toyota / Woven by Toyota 软件平台
- **TPS** — Toyota Production System (丰田生产方式)
- **HIMA** — Huawei Intelligent Mobility Alliance (华为智选车)
- **ADS** — Autonomous Driving Solution (华为自动驾驶方案)
- **MDC** — Mobile Data Center (华为车载计算平台)
- **XNGP** — Xpeng Navigation Guided Pilot (小鹏导航辅助驾驶)
- **VLA 2.0** — 小鹏 Vision-Language-Action 第二代
- **Turing chip** — 小鹏自研 AD/ADAS 推理芯片
- **HAD** — Xiaomi Autonomous Driving (小米自研自动驾驶)
- **评估方法学**
- **Tier 1-5** — 数据来源权威性金字塔层级
  - Tier 1 = SEC 文件 / IPO 招股书 / 公司财报 / 股东大会披露
  - Tier 2 = 财报电话会议 / 官方技术发布会
  - Tier 3 = 行业分析师 / 同行评审论文
  - Tier 4 = 主流技术媒体报道
  - Tier 5 = 社交媒体非正式发言 (如 Musk X 帖)
- **可证伪点** — 可被未来工程现实实证或反证的具体场景, 本报告第 10.6 节列出 5 个

**数据核实纪律声明:** 本报告涉及 Tesla 硅规格、FSD 版本、Optimus 规格、训练基础设施, 所有关键技术参数均经 2026 年 5 月公开来源核实。对马斯克口径数据(如”40x faster”)标注口径性质; 对来源不一的数据(如 Optimus 全身自由度、Cortex 集群规模)标注差异区间而非强行统一。本报告 Tier 1 来源包括 Tesla Q1 2026 Update (assets-ir.tesla.com)、Tesla 10-Q / 8-K (SEC.gov)、SpaceX S-1 招股书 (2026.5.20)、Tesla Annual Shareholder Meeting (2025.11.6)、Xpeng SEC Form 6-K、Xiaomi HKEX 年报、VW Group China Investor Update 等。

---

## 摘要

作为 archi-intelligence Research Series 的第三篇、亦是收官之作,本研究以 Tesla FSD-Optimus 统一技术栈为实证锚点,闭合 D1(《架构的世纪迁徙》,2026-01)所提出的 AR0-AR5 架构能力门槛框架与 D2(《2026 全球汽车 E/E 架构成熟度评估》,2026-02)对 22 家 OEM 的横向成熟度评估。截至 2026 年 5 月,Tesla 是全球唯一同时满足 AR4(多形态物理 AI 平台)全部判据并已进入量产爬坡的工程实体;本研究选取它并非为其商业前景背书,而是基于一个方法论判断——它是当前唯一一个将官方披露的 11 层共享核心技术真正打通于汽车与人形机器人两种形态的案例。通过在硅路线(HW3 → AI7,含 AI4.5 stopgap 与 AI5“不给车”的硬件分级)、FSD 软件栈(从 V11 规则到 V14 端到端 + VLA 的 Software 2.0 范式革命)、形态无关的占用网络、训练基础设施(Dojo → Cortex 2.0,以及 SpaceX-xAI 合并后的集团级算力重构)等层面的临床剖析,本研究将 Tesla 官方披露的 11 层共享技术提炼为三组结构迥异的复用机制——物理层(6 项)、智能层(5 项)与一个隐含的组织层(第 12 层,即 FSD-Optimus 团队合并)——并精确刻画跨形态复用的边界:积累性能力可复用,而形态依赖的高频运动控制不可复用。本研究的核心论点是,Tesla 是一个 **AR4 参照系**而非可复制的模板:没有任何追赶者能复制这条路径所要求的 20 年积累、全栈垂直与集团级资源配置,但每一家追赶者都能借此参照系精确定位自身的真实位置与合理路径——这一判断通过对五条不同轨迹(大众、丰田、华为、小鹏、小米)的 12 层深度镜像得到具体论证。在实证层面,Tesla 案例支撑了 D1 的两层结构论、“AI 是语言能力而非大脑”命题(由 Tesla 自身的部署节奏与 SpaceX S-1 中对 scaling law 的风险提示双重印证)、架构债理论与失效哲学框架,同时本研究亦列出五个可证伪点,供 D1 框架在 2027-2028 年的工程现实中接受检验。本研究最终指出一个结构性市场空白:当工业工具链已为详细设计阶段提供成熟支撑时,概念探索阶段——在画第一张图、写第一行代码之前做出关键架构决策的阶段——仍依赖个人与组织的工程直觉,而这正是追赶者没有 20 年时间去积累的能力。

**关键词:** AR4;多形态物理 AI 平台;跨形态复用;垂直闭环;Tesla FSD;Optimus;占用网络;Software 2.0;架构就绪度;失效哲学;架构智能

## 引言：为什么 Tesla 是 AR4 垂直闭环的唯一完整案例

### 本研究在 archi-intelligence Research Series 中的位置

本报告是 archi-intelligence Research Series 的第三篇 Working Paper（2026-03）。理解 D3 的价值，必须从其在系列中的三重位置理解：

**第一重——D3 是 D1 理论框架的实证锚点。**D1（《架构的世纪迁徙》，2026-01）提出了 AR0-AR5 架构能力门槛框架、AI<sup>2</sup>-ML 架构智能成熟度阶梯、两层结构论（基础设施收敛 vs 控制语义发散）、失效哲学（fail-soft vs fail-operational）等核心判断。这些判断作为理论框架，需要工程实证的检验。**D3 选取 Tesla 作为” AR4 唯一完整闭环” 案例，将 D1 的理论判断置于具体工程颗粒度下做临床检验——哪些得到支撑、哪些受到挑战、D1 框架的可证伪点在哪里。**第十章是这一检验的集中呈现。

**第二重——D3 是 D2 全景评估的深度延伸。**D2（《2026 全球汽车 E/E 架构成熟度评估》，2026-02）对 22 家 OEM 做了横向评估，给出 Snapshot 与 Roadmap 双时间维度评分。但横向评估的局限性是：**它能给出每家 OEM 的位置，却难以揭示” 为什么位置在这里、向上跃迁的路径是什么”**。D3 通过对 D2 中唯一的 AR4 双雄之一（Tesla，另一为华为 HIMA）做深度切片，**为 D2 评估中的其他 21 家 OEM 提供了一个 AR4 参照系。**第九章是这一镜像价值的具体展开——大众、丰田、华为、小鹏、小米五条不同的” AR4 企图路径”，在 Tesla 工程颗粒度的参照下，其各自的位置、约束、合理路径变得清晰可见。

**第三重——D3 揭示了 AR3+ 向 AR4 跃迁的市场空白。**当 D1 提供理论、D2 提供评估、D3 提供工程颗粒度时，三者共同回答了一个核心问题：**当其他 21 家 OEM 不可能复制 Tesla 时，他们向上跃迁的合理路径需要什么样的方法论与工具支撑？**第十章第 10.5 节通过对追赶者核心矛盾的层层推导，指出当前工业工具生态对**详细设计阶段**已有成熟支撑（EDA、PLM、安全证明工具链等），但对**概念探索阶段**——架构师在画第一张图、写第一行代码之前需要做出关键架构决策的那一阶段——仍主要依赖个人与组织级的工程直觉。Tesla 通过 20 年累积内化了这一能力；追赶者没有 20 年时间。这构成了一个值得产业界、学术界、工具供应商共同讨论的市场空白。

### Tesla 选题的方法论判断

在 D1 旗舰报告中，我们提出了 AR0-AR5 架构能力门槛框架，并将 AR4（多体物理 AI 平台）定义为这样一个阶段：训练、仿真、部署形成统一闭环；世界模型与基础模型跨车辆与机器人共享；模型复用率高；策略迁移效率高（跨形态泛化）。

截至 2026 年初，全球范围内，**唯一同时满足 AR4 全部判据并已进入量产爬坡的案例，是 Tesla 的 FSD-Optimus 统一栈。**这并非对 Tesla 商业前景或安全争议的背书——本报告对其纯视觉路线的监管风险、责任归属脆弱性、CEO 个人风险均保持批判性审视。我们选择 Tesla 作为 D3 深度解剖对象，纯粹基于一个方法论判断：**Tesla 是当前唯一一个将官方披露的 11 层共享核心技术（执行器、功率电子、电池、制造、数据通信、音频、摄像头、AI 芯片、训练集群、神经仿真、现实世界 AI）跨汽车与机器人形态完整打通，且通过组织层整合使这一跨形态复用真正成立的工程实体。**

理解 Tesla 这一案例的价值，不在于”学习如何成为 Tesla”——正如旗舰报告所论证，垂直闭环要求的资本规模、人才密度、组织扁平度是 20 年累积的产物，不可在 18 个月内复制。其价值在于：**Tesla 提供了一个 AR4 架构的”参照系”，让我们能够精确刻画”跨形态复用到底在工程上意味着什么、哪些可复用、哪些不可复用、代价是什么”**。

### 本报告的研究边界

本报告明确承认三重限制：

第一，**公开度限制**。Tesla 通过 AI Day（2021、2022）、各季度财报、马斯克 X 平台发言、最新的 SEC 10-Q/8-K 文件，以及 SpaceX S-1 招股书（2026.5.20）提供了相对丰富的技术披露，但仍有大量底层实现细节（如 FSD 端到端网络的精确架构、Cortex 2.0 的具体拓扑、AI5 的微架构）未公开。本报告严格区分”已公开事实”（标注 Tier 1-2 来源）与”基于公开信息的方向性推断”。

第二，**马斯克口径问题**。Tesla 的许多性能声明来自马斯克的公开发言，其数字（如”AI5 比 AI4 快 40 倍”）往往是综合性、营销性的口径，与标准化的 TOPS、FLOPS 度量不直接可比。本报告对这类数据标注其来源性质，不作为客观比较基准采用。

第三，**时间窗口限制**。本报告的数据截止至 2026 年 5 月 21 日。Tesla 案例处于高速演进期，本报告所述事实可能在数月内被新事件刷新。我们承诺在后续版本中公开记录关键事实的更新与修正。

## 本报告的结构

本报告共分十章，加引言与结论：

- **第一章 硅路线**：从 HW3 到 AI7 的代际演进，含 AI4.5 stopgap、AI5 不给车的”硬件分级”、Terafab 集团级算力战略
- **第二章 FSD 软件栈**：从 V11 模块化规则栈到 V14.3 端到端 + VLA 多模态的范式革命，含 Robotaxi 7 城承诺改口与中欧监管现实
- **第三章 占用网络**：作为跨形态复用的形态无关 3D 表示
- **第四章 跨形态复用机理**：Tesla 官方 11 层共享核心技术 + 组织层第 12 层的三组结构提炼（物理 / 智能 / 组织）
- **第五章 Optimus 运动学与不可复用层**：差异的工程颗粒度
- **第六章 训练基础设施**：Dojo → Cortex → COLOSSUS/COLOSSUS II 的演化（含 SpaceX-xAI 合并后的归属重构）
- **第七章 组织层面的复用**：FSD/Optimus 团队合并 + Musk 跨 8 实体控制的极端组织债务
- **第八章 AR4 的代价与边界**：对其他玩家的启示，含 SpaceX S-1 AI capex 硬数据
- **第九章 镜像章**：D3 对 D2 其他 OEM 的论证意义（大众、丰田、华为、小鹏、小米深度镜像）
- **第十章 实证章**：D3 对 D1 核心观点的支撑与挑战，含工具范畴反向定义 + D1 可证伪点
- **结论**：AR4 作为参照系的方法论意义

每章都明确标注其在 archi-intelligence Research Series 三重价值中的位置（理论实证、评估延伸、追赶者工具范畴推导），以便读者建立 D1 → D2 → D3 三部曲的闭环理解。

## 第一章 硅路线：从 HW3 到 AI7 的代际演进

Tesla 自动驾驶硬件的演进，是理解其 AR4 架构的物理起点。与传统 OEM” 采购 Tier-1 提供的 SoC” 不同，Tesla 自 HW3 起即走上自研推理芯片的道路。这一决策的深远影响在于：**芯片、感知算法、训练基础设施由同一组织在同一设计意图下协同演进**，从而消除了传统供应链中” 芯片厂商不懂算法、算法团队迁就芯片” 的错配。然而，截至 2026 年 5 月，这条路径已开始显露其工程边界：AI5 的延期产生了 AI4.5 这一 stopgap 产物，AI5 首批硅片优先供给 Optimus 与超算集群而非车端，存量 HW3 车队被官方确认无法达到无监督 FSD 的能力门槛。这些事实，恰恰是 AR4 路径不可绕过的代价。

### 1.1 代际总览

表 1.1: Tesla 自动驾驶硅代际对照（截至 2026 年 5 月）

世代	商品名	算力	制程	代工	状态	内存
HW3	FSD Computer	~144 TOPS	14nm	Samsung	存量主力；官方确认无法跑 <b>unsupervised FSD</b>	8GB LPDDR4
HW4	AI4	~500 TOPS	Samsung 7nm	Samsung	新车主力，Cybercab Q2 2026 SOP 仍用此 <b>2026 Model Y 起悄然搭载</b> （因 AI5 延期产生的 stopgap）	16GB LPDDR5
—	AI4.5	AI4 的精细化升级	Samsung 7nm	Samsung		—
HW5	AI5	~4,000 TOPS 等效（5x AI4 useful compute / 8x raw）	先进节点（TSMC/Samsung 双源）	TSMC Arizona + Samsung Taylor TX	<b>2026.4.15 流片</b> ；engineering samples 2026 末； <b>量产 mid-2027；首批给 Optimus + super-computer，不给车</b>	192GB LPDDR5X

世代	商品名	算力	制程	代工	状态	内存
—	AI6	进一步演进	同 AI5	同 AI5	目标 2026.12 流片	—
—	AI7	待定	待定	待定	早期规划	—

关键事实核实（来源：Tesla Q1 2026 Update PDF、SEC 10-Q FY2026 Q1、Q1 2026 财报电话会议、SpaceX S-1 招股书 2026.5.20）：

- AI5 较 AI4 的官方提升口径：~5× useful compute、8× raw compute、9× on-chip memory (16GB→192GB)、5× memory bandwidth。
- 马斯克”AI5 比 AI4 快 40 倍”为综合性营销口径（含带宽、架构效率、内存等组合），非纯算力倍数。本报告标注此口径性质，不将其作为客观比较基准。
- AI5 高量产推迟到 mid-2027——这是与马斯克 2024.6 原承诺（“AI5 装车 2H 2025”）相比的近两年延期。
- Cybercab（2026.2.17 首台下线，Q2 2026 SOP）确认使用 AI4，而非原计划的 AI5。
- HW3 unsupervised FSD 已被官方放弃（Q1 2026 财报电话会议，Musk 与 Ashok Elluswamy 共同确认）。Tesla 提供 HW3 车主两条路径：折价换购 AI4 车型，或免费 computer+camera 改装。HW3 的 v14 Lite 版本承诺 2026.6 前推送。

## 1.2 HW3：自研推理芯片的起点（2019）与无监督承诺的撤回

HW3（FSD Computer）于 2019 年量产，是 Tesla 首款完全自研的自动驾驶推理芯片。其约 144 TOPS 的算力在当时已显著领先于行业，但更重要的是其架构哲学的确立：双 NPU 冗余设计（两颗独立的神经网络加速器互为校验）、为视觉推理工作负载专门优化、与 Tesla 自研的视觉算法栈协同设计。

然而，2026 年 4 月 22 日的 Q1 2026 财报电话会议正式确认：HW3 无法达到无监督 FSD 的能力门槛。这一确认本身具有方法论意义——它揭示了 AR4 路径的一个内生矛盾：当公司试图将日益庞大的端到端神经网络（FSD V12 起）压缩进 2019 年的 HW3 算力包络时，硬件代际的物理上限成为软件演进的真实约束。Tesla 自 2019 年起就向 HW3 用户出售（FSD 包售价 \$8,000-\$15,000）“硬件足够支持完全自动驾驶”的承诺；这一承诺在 2026 年的撤回，是垂直闭环企业在”对存量用户的承诺”与”硬件物理极限”之间张力的典型体现。

补偿机制本身揭示了垂直整合的代价：Tesla 必须自行承担 HW3 车队的改装成本（约百万辆车，自付硬件 + camera retrofit + 城市级 microfactory 网络）——这一成本若在传统”硬件采购”模式下，可被部分转嫁给 Tier-1 供应商；但在 Tesla 的垂直闭环下，由企业自身完整承担。

## 1.3 HW4 / AI4：当前量产主力（2023）与 AI4.5 的 stopgap 浮现

HW4（AI4）于 2023 年装车 Model S/X refresh，2024 年起逐步覆盖 Model 3/Y/Cybertruck。其约 500 TOPS 算力相对 HW3 提升约 3.5 倍，16GB LPDDR5 内存允许更大的端到端神经网络部署。

值得注意的事实是 AI4.5 的悄然出现：因 AI5 延期，Tesla 于 2025 年末在 2026 Model Y 中引入了 AI4.5——一颗 AI4 的精细化升级版（核心硅设计未变，但工艺与内存配置有调整），用于在 AI5 高量产之前维持 FSD 神经网络的算力需求。Tesla 未对 AI4.5 做高调发布；其存在本身揭示了一个事实：AI5 的延期幅度，已经超出了 AI4 单代足以支撑的窗口。

Cybercab——Tesla 的专用 Robotaxi 车型——2026 年 2 月 17 日首台下线，Q2 2026 SOP，初期采用 AI4 而非 AI5。这一决策在 Q1 2026 财报中明确：Tesla 此前的承诺是 Cybercab 用 AI5；现实是

Cybercab 必须在 AI4/AI4.5 平台上启动量产。这意味着 **Cybercab 早期车队的算力上限将受 AI4 包络约束**，而非按马斯克长期叙事所述的 AI5 跃升。

#### 1.4 AI5 / HW5: 跨代际跃升 (2026.4 流片) 与”不给车”的硬件分级

2026 年 4 月 15 日，马斯克在 X 平台公布 AI5 流片照片，宣布 Tesla 完成对 TSMC 的 GDSII 数据交付。流片即芯片设计的最终定稿；从流片到量产仍需 12-18 个月。

AI5 的工程跃升幅度，远超 HW3→HW4:

- **算力等效约 4,000 TOPS** (5× AI4 useful compute / 8× raw / 9× on-chip memory)
- **内存大幅扩展到 192GB LPDDR5X** (HW4 的 16GB → AI5 的 192GB, 扩大 12 倍)
- **去除 ISP / 图形单元**——AI5 设计为纯推理 GPU，连 image signal processor 都被设计团队删除。这意味着 AI5 不会作为传统 SoC 出售给其他车企；它是 Tesla 全栈推理工作负载的专用芯片。
- **双代工战略**: TSMC Arizona + Samsung Taylor, TX——皆位于美国本土，规避关税风险并实现供应链冗余。

**但 AI5 的硬件分级是 D3 第一章必须突出的”打脸点”：**

Q1 2026 财报电话会议中，马斯克直接确认——**AI5 首批硅片，将优先供给 Optimus 人形机器人与 Tesla / xAI 的超算训练集群，而非车端**。Cybercab Q2 2026 SOP 与 2026 全年新车，都将继续使用 AI4 / AI4.5。AI5 的车端高量产，延迟到 mid-2027。

这一事实，与 Tesla 此前的叙事形成明显反差。2024 年 6 月，马斯克曾承诺”AI5 装车 2H 2025”；2025 年 11 月，时间表推迟到”2026 末-2027 初”；2026 年 4 月流片时，Cybercab 已被官方确认仍用 AI4。换言之，当车端原本是 AI5 设计意图的核心使用场景时，工程现实是 **Tesla 自己的 Optimus 与 xAI 集群对 AI5 算力的优先级，高于车端推理**。

这并非偶然，而是 Tesla AR4 闭环的内生结构——当同一推理芯片必须服务 (1) 车端实时推理、(2) Optimus 端边缘推理、(3) 超算训练任务时，资源稀缺会自然向最高价值工作负载倾斜。Optimus 与 xAI 集群处于产品定义的”最贵阶段” (小批量、高单价、高战略意义)；车端处于”成熟摊销阶段” (大批量、低单车毛利贡献边际、AI4 已足够支撑 supervised FSD)。Tesla 用 AI5 给 Optimus / xAI 不给车，并非工程失败，而是 AR4 垂直闭环下资源配置的逻辑必然。

这一点，恰恰强化了 D3 第四章 4.3 节智能层论证的论点：**Tesla 的”统一栈”不是同一颗芯片跑两个形态，而是同一软件栈跨两种算力等级 (车端 AI4/AI4.5 + 机器人/超算 AI5)**，通过软件抽象层吸收硬件差异。

#### 1.5 AI6 / AI7: 未来路标与 Terafab 计划

Q1 2026 财报电话会议中，马斯克进一步披露：

- **AI6 流片目标 2026 年 12 月**——若实现，将代表 9-12 个月的”代际节奏”，远快于半导体行业典型的 18-24 个月架构升级周期。
- **AI7 已在早期规划**。
- **Dojo 3 芯片项目”与 AI6 同步推进”**。

但更具结构意义的事实是 **Terafab 计划**——由 SpaceX 与 Tesla 于 **2026 年 3 月**联合公布、Intel 于 **2026 年 4 月**加入的”芯片制造创新项目” (投资规模约 \$25B, 选址 Austin, Texas)。SpaceX S-1 招股书 (2026.5.20) 给出了 Terafab 的官方定义：

“Terafab — a chip manufacturing initiative with a long-term goal of producing one terawatt of compute hardware each year.” (每年生产一万亿瓦特算力硬件的芯片制造

倡议)

S-1 同时明确: Terafab 当前仅是“**general framework**”, “具体项目仍需单独协商与协议”(including any development timelines, milestones and capital expenditures), 尚未敲定具体出资比例与项目时间表。

Terafab 的战略意图, 在 SpaceX S-1 的” AI Compute Infrastructure” 章节有清晰陈述:

“We believe that the key constraints in the continued growth of AI are physical—chip manufacturing, data center infrastructure, and power generation; the future of AI will be determined by the control of the physical stack.” (我们相信 AI 持续增长的关键约束是物理性的——芯片制造、数据中心基础设施、电力生成; AI 的未来将由对物理栈的控制权决定。)

这段陈述揭示了 Terafab 的真正用户: **SpaceX 自身的 orbital AI compute satellites 部署计划** (2028 年起开始部署, 目标 100GW/年算力发射至 Sun-synchronous orbit)。Tesla 是 Terafab 框架的协作方, 但 Tesla 车端 AI 芯片的采购量, 远小于 SpaceX 轨道 AI 卫星所需的”百万吨级”硬件部署。

**对 D3 的方法论意义:** Tesla 自研芯片路径, 已经从”汽车业务的垂直闭环组成部分”演化为”SpaceX-xAI-Tesla 集团级 AI 算力战略的下游”。这是组织债务的极端形态——当马斯克同时控制 Tesla、SpaceX、xAI、X、Neuralink、Boring、Macrohard、Terafab 8 家实体的董事会多数席位时, 资源配置不再服从单一公司的最优解, 而服从马斯克个人视角下的”集团最优解”。**第七章将进一步展开这一组织复用的结构含义。**

## 1.6 硅路线的架构启示

Tesla 硅路线给出四个对 AR4 架构的关键启示:

第一, **自研推理芯片是 AR4 的硬件前提, 但非充分条件。**NVIDIA、华为、Mobileye 也有强大的自研芯片, 但 Tesla 的独特之处在于芯片设计意图与上层算法、训练基础设施的同源协同——这是组织层面的整合, 而非单纯的芯片能力。AI5 去除 ISP/图形单元演化为纯推理 GPU 的决策, 恰恰是这一同源协同的工程结果——传统 SoC 厂商 (即使技术能力强如 NVIDIA) 无法做出这种针对性专业化, 因为它们必须服务多个客户的多样化工作负载。

第二, **硬件代际与软件演进的张力, 是垂直闭环的代价而非副作用。**HW3 unsupervised FSD 撤回事件揭示了一个普遍规律: 当软件演进速度超过硬件代际节奏时, 对存量用户的承诺成为工程负担。这一矛盾在”硬件采购”模式下被部分转嫁给供应商, 但在垂直闭环下由企业自身完整承担。Tesla 的 microfactory 改装网络与免费 retrofit, 是这一代价的具象化。

第三, **AR4 路径必然产生”硬件分级”。**AI5 首批不给车而给 Optimus / xAI 集群的事实, 看似与 AR4”统一栈”叙事矛盾, 实则揭示了 AR4 闭环的真实结构——**不是同一颗硅片服务所有形态, 而是同一软件栈跨多种算力等级运行, 由软件抽象层吸收硬件差异。**这一区分, 对追赶者的工具选择具有方向性意义: 若试图通过”采购一颗强算力芯片”复刻 Tesla 的 AR4, 本质上是误读了 AR4 的真正机理。

第四, **硅路线已超出汽车业务边界, 演化为集团级 AI 算力战略。**Terafab 框架表明, Tesla 自研芯片的下游用户已包括 SpaceX 的轨道 AI 算力卫星——这是 Musk 集团级资源配置的明确信号。对其他 OEM 而言, 这意味着: **追赶 Tesla 的 AR4 路径, 不仅需要追赶其汽车业务的垂直闭环, 还隐含追赶一个跨越汽车-机器人-太空的集团级资源分配能力。**这是不可复制的。



图 1.1 Tesla 硅代际演进时间线 — From HW3 to AI7: 从汽车业务的垂直闭环演化为集团级 AI 算力战略。三个红色★“标记为关键打脸点：HW3 unsupervised FSD 撤回（2026.4）、AI4.5 stopgap chip（因 AI5 延期产生）、AI5 首批不给车（优先 Optimus + 超算）。Terafab 集团级算力战略横幅作为上层语境。

## 第二章 FSD 软件栈演进：从规则到端到端的范式革命

如果说硅路线是 Tesla AR4 架构的物理起点，那么 FSD 软件栈的演进则是其灵魂。从 V11 到 V14，Tesla 完成了一场自动驾驶领域罕见的**范式革命**——从手写规则驱动转向端到端神经网络驱动。这一转变不仅是技术路线的选择，更是 AR4”训练-仿真-部署统一闭环”判据得以成立的软件前提。截至 2026 年 5 月，V14 已迭代至 v14.3.3，并出现了一个值得方法论关注的事实：**马斯克在 Q1 2026 财报中明确表示，“FSD v14.3 软件栈在架构上已足以支持 unsupervised 部署，剩下的是 validation 与监管审批，不是技术能力问题”**。这一陈述为 D3 第十章关于”AI 是语言能力非大脑”的论证提供了关键素材——若 Tesla 自己的判断正确，则验证与监管是 AR4 的真实瓶颈，而非神经网络能力本身。

### 2.1 V11 及之前：模块化规则栈的黄昏

在 V12 之前，Tesla FSD 采用的是行业主流的**模块化架构**：感知（perception）、规划（planning）、控制（control）三个模块通过明确定义的接口交互。感知模块输出边界框（bounding box）、车道线段等结构化结果，规划模块基于这些结果做决策，控制模块执行。

这一架构的核心瓶颈在于**接口信息损失**。传统方法中，感知结果以边界框或线段形式传递，这限制了信息的表达力——真实世界的复杂场景被压缩成有限的结构化标签，大量隐含语义在模块间传递时丢失。V11 是这一模块化范式的集大成者，也是其黄昏。据公开信息，V11 包含约 30 万行以上的 C++ 控制代码——这些手写规则覆盖了无数特定场景，但也成为难以维护、难以泛化的技术债务。

值得注意的是，**V11 至今仍在 HW3 车架上运行**（具体为 v12.6 分支），与 HW4 车队的 V14 分支并行存在。Tesla 已承诺为 HW3 车队推出”V14 Lite”——一个针对 HW3 算力包络的精简端到端版本——但这一承诺在 Q1 2026 财报中再次明确”by end of June 2026”，意味着已比原计划晚约 6 个月。这一软件版本分叉，本身是 AR4 路径下”硬件代际与软件演进张力”的另一具象。

## 2.2 V12: 端到端革命 (2024.1)

V12 是 Tesla FSD 历史上最重大的架构断裂。其核心变革是：**用单一端到端神经网络替代约 30 万行 C++ 控制代码**。Tesla 将这一范式概括为” Photon In, Control Out”（光子进，控制出）——从摄像头像素输入到车辆控制输出的整个过程，由单一神经网络完成，不再有显式的感知/规划/控制模块边界。

这一转变的本质，是将自动驾驶问题重新定义为**模仿学习问题**。端到端大模型本质上将海量驾驶视频数据集压缩进网络参数，这与生成式语言模型将互联网规模文本压缩进模型参数高度类似。FSD 因此成为”软件 2.0”时代的产物——完全由数据驱动，而非由工程师手写逻辑驱动。

V12 解决了模块化系统的”接口瓶颈”：感知与规划/控制不再通过有损接口传递信息，而是合并为单一神经网络结构，信息在网络内部以高维特征形式流动，避免了结构化标签的信息压缩损失。

## 2.3 V13: HW4 原生分辨率 (2024.12)

V13 是 V12 范式在 HW4 硬件上的深化。其关键改进包括：HW4 原生分辨率输入（充分利用 AI4 更高的摄像头带宽与算力），以及训练数据规模相较 V12 的约 4.2 倍扩展。V13 证明了端到端范式的**数据可扩展性**——更多数据、更高分辨率、更大算力直接转化为能力提升，这正是”软件 2.0”范式的核心承诺。

## 2.4 V14: 多模态与世界模型 (2025-2026)

V14 是当前 (2026.5) 的主力版本。截至 2026 年 5 月中旬，公开追踪的最新版本为 **V14.3.3 (固件 2026.14.6.6)**，于 2026.5.17 开始向早期访问用户推送，主要新增 **“intervention-free live streak counter”**——一个实时显示”自上次人工干预以来连续行驶里程”的计数器。Tesla 的官方意图，是通过让用户在车端看到”无干预连续里程”的累计，建立 unsupervised FSD 即将到来的市场预期。

V14 的核心网络呈现出显著的**多模态化特征**：

**输入 (多模态)**：- 7 路高分辨率摄像头视频 - 车辆自身运动信息 - 导航信号 - 音频信号

**输出 (多任务)**：- 语义分割 (semantic segmentation) - 占用栅格 (occupancy grid) - 3D 高斯特征 (3D Gaussian features) - 语言表达 (language expression) - 最终控制动作 (control action)

这一输入输出结构揭示了一个关键趋势：**FSD V14 可能已接入视觉-语言-动作 (VLA) 框架**，赋予模型”解释”与”思考”的能力。语言表达输出的存在，意味着模型不仅在做”感知-控制”映射，还在构建对场景的语义理解——这与中国新势力 (小鹏 VLA 2.0、理想 Mind GPT VLA、华为 ADS 4.0 WEWA) 的演进方向高度收敛。

马斯克曾透露，开发中的 FSD V14 系统参数规模相较 V13 扩展约 4.5 倍，这一参数膨胀正是 AI5 高算力需求的来源——HW4 的算力包络已难以承载 V14 的完整潜力。但在 AI5 高量产推迟到 mid-2027 的现实下，V14 的车端部署必须在 AI4/AI4.5 上完成——这一约束反过来塑造了 V14 的工程优化方向，包括 MLIR-based AI 编译器的重写 (v14.3 release notes 提及) 以最大化既有硬件算力利用率。

**数据性质标注**：部分第三方分析将 V14 描述为采用”生成式世界模型 (GWM) “、将环境视为”4D 仿真”。这些表述源于第三方博客解读，Tesla 官方未明确确认”GWM”术语。

本报告将其标注为方向性解读，不作为 Tesla 官方架构定义采用。

## 2.5 Robotaxi 部署与监管现实

V14 的能力进展，必须与 Tesla Robotaxi 商业化部署的实际进度对照评估。

**部署现状**（截至 2026 年 5 月）：- 2025 年 6 月，Robotaxi 在 Austin 启动（Tesla 首个 unsupervised 商业运营城市）- 2026 年 4 月 18 日，扩展至 Houston 与 Dallas - 当前共 **3 个城市**（Austin / Houston / Dallas），均在 Texas 州 - 独立观察者（基于车牌追踪）指出，每个城市的实际运营车辆数仅为 **2 辆左右**——这是一个极小的样本量

**承诺改口**：Tesla 此前（2025 Q4 财报）承诺 1H 2026 内运营 7 个美国城市——Austin、Houston、Phoenix、Miami、Orlando、Tampa、Las Vegas。Q1 2026 财报中，**剩余 5 城的措辞从”1H 2026”改为模糊的”preparations underway”**。距 1H 2026 仅剩 2 个月，这一改口实质构成对 5 城承诺的延期。

**国际监管**：- **中国**：Q1 2026 财报确认 supervised FSD 在中国已开始向用户推送，full regulatory approval 预期 Q3 2026。Tesla 通过 2026 年初的 OTA 更新（version 2026.2.9）启动了对中国市场的本地化。- **欧洲**：2026 年 4 月，Tesla 与荷兰 RDW（Netherlands Vehicle Authority）“已完成 supervised FSD 最终车辆测试阶段”。FSD V14 在欧洲的实际部署需通过 UN-R157 等多重法规审核。

Robotaxi 部署的现实节奏，与 Musk 在 2024-2025 年提出的”unsupervised FSD 即将解决”叙事形成显著差距。**这一差距本身是 D3 第十章”AI 是语言能力非大脑”论证的关键素材**——若 FSD V14.3 已”在架构上足以支持 unsupervised”（马斯克 2026.4 公开陈述），那么剩余的”validation + 监管审批 + 城市级 deployment”工作量，已超过单纯的神经网络能力构建本身。这印证了 D1 旗舰报告的判断：**AI 是语言能力，骨架是确定性的安全证明 / 验证 / 监管协同**。

## 2.6 软件栈演进的架构启示

FSD 软件栈的演进对 AR4 架构有四个关键启示：

第一，**端到端范式是跨形态复用的算法前提**。模块化架构中，针对汽车手写的规划规则无法迁移到机器人；但端到端神经网络学到的是”从感知到行动”的通用映射能力，这种能力在原理上可跨形态迁移（详见第四章）。

第二，**软件 2.0 范式将能力提升转化为数据与算力问题**。一旦架构从”手写规则”转为”数据驱动”，能力的天花板就从”工程师能想到多少场景”变为”能采集和训练多少数据”。这正是 Tesla 数据闭环（C5 控制点）价值的根源。

第三，**多模态 VLA 化是跨产业收敛的信号**。FSD V14 接入 VLA 框架，与中国新势力（小鹏、理想、华为）的演进方向一致，印证了旗舰报告”AI 模型架构跨域通用”的判断——Transformer + 占用 + VLA 正成为物理 AI 的通用架构基底。

第四，**软件能力 ≠ 部署能力**。Tesla 自己（2026.4 财报）确认 FSD v14.3 ”架构上足以 unsupervised”，但 Robotaxi 实际部署仅在 3 城 × 每城 2 辆运营，且 7 城承诺已悄然改口。这一差距揭示了 AR4 落地的真实瓶颈——**不是神经网络能力，而是验证、安全证明、监管协同、城市级运营基础设施**。这是 D1 ”AI 是语言能力，骨架是确定性引擎”论点的临床实证（详见第十章）。

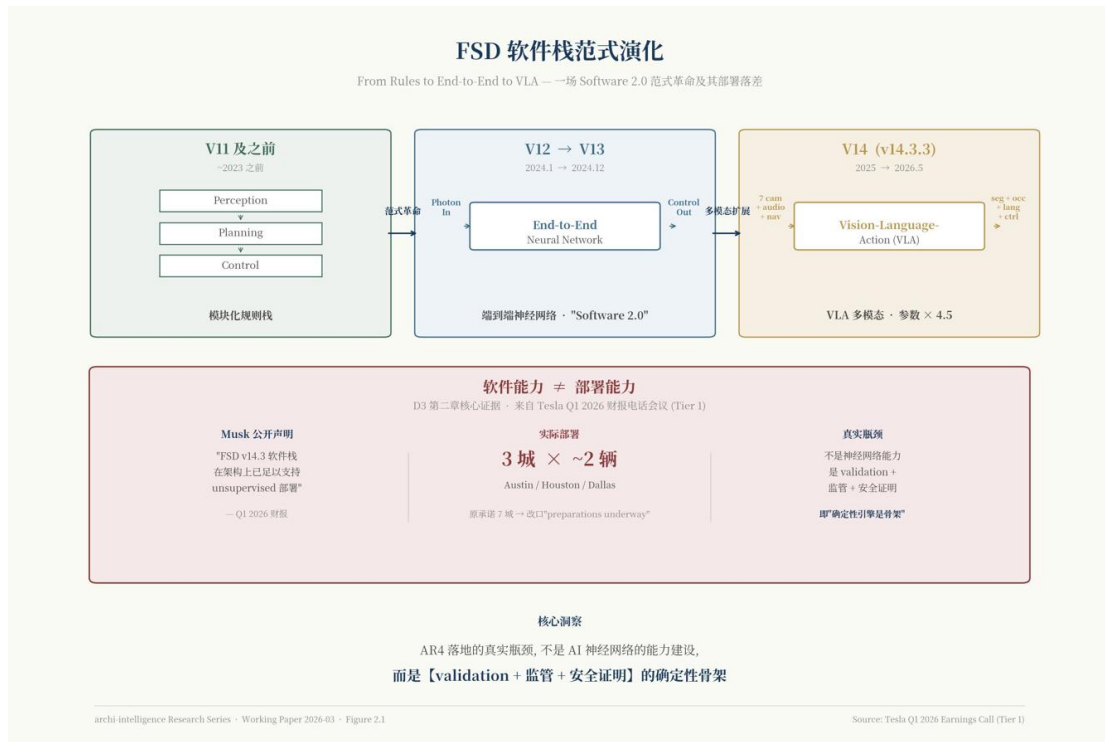


图 2.1 FSD 软件栈范式演化 — From Rules to End-to-End to VLA: 一场 Software 2.0 范式革命及其部署落差。三段范式 (V11 模块化规则栈 → V12-V13 端到端 → V14 VLA 多模态) 展示软件能力跃迁; 中央红色对比块显示 Musk 公开声明 (“v14.3 架构上足以 unsupervised”) vs 实际部署 (3 城 × ~2 辆) vs 真实瓶颈 (validation + 监管 + 安全证明) 的三重落差。

### 第三章 占用网络与形态无关感知

如果说端到端范式是跨形态复用的”算法前提”，那么**占用网络 (Occupancy Networks)** 就是使这一复用在感知层得以实现的”核心技术机制”。理解占用网络，是理解 FSD 栈为何能迁移到 Optimus 的关键。

#### 3.1 占用网络的技术本质

占用网络由 Tesla 在 2022 AI Day 公开。其核心思想是：**将世界表示为一个体素化 (voxelized) 的 3D 稠密占用场**——空间被划分为大量小立方体 (体素)，网络预测每个体素是”被占用”还是”空闲”，并叠加”占用流 (occupancy flow)” “预测每个被占用体素的运动方向。

这一表示方法与传统的”边界框 + 分类”感知有本质区别：

- **边界框感知**：将物体识别为预定义类别 (车、人、自行车) 的矩形框。无法表达”未知形状的障碍物”——一个翻倒的货物、一团不规则的碎片，在边界框范式下难以处理。
- **占用网络**：不关心物体”是什么”，只关心空间”是否被占据、如何运动”。这使得系统能处理任意形状的障碍物，无需预先定义类别。

#### 3.2 为何占用网络是”形态无关”的

占用网络的革命性，在于它产生的是**形态无关 (embodiment-agnostic) 的 3D 场景表示**。

考虑这一表示的本质：占用网络从相机像素生成的，是关于”周围 3D 空间哪里被占据、如何运动”的稠密理解。这一理解**不依赖于感知主体是什么形态**——无论搭载相机的是一辆汽车、一个人形机器人、还是一架无人机，“周围空间的占用状态”这一物理事实是客观的、形态无关的。

正是这一属性，使得 FSD 的视觉感知栈可以直接迁移到 Optimus。Tesla 公开表示，得益于占用网络的形态无关性，FSD 栈迁移到 Optimus 时**体素大小（voxel size）可降至 10cm**——以适应机器人在室内精细操作（如抓取、避障）对空间分辨率的更高要求。汽车场景下 10cm 体素或许偏精细，但同一套占用网络架构通过调整体素粒度即可适配机器人场景，这正是“同源感知栈”的工程体现。

### 3.3 占用网络在 Tesla 感知栈中的位置

占用网络并非孤立存在，而是嵌入在 Tesla 的整体感知架构中。其演进脉络为：

早期 Tesla 使用 HydraNet（多头神经网络）做多任务感知——一个共享主干（backbone）+ 多个任务头（检测、分割、车道线等）。占用网络引入后，HydraNet 被扩展，占用网络的输出被规划器（planner）整合，提升了 3D 场景理解能力。在 V12 端到端化之后，占用表示进一步融入单一端到端网络的内部特征流——如 V14 的输出中仍包含“占用栅格”这一项，说明占用表示作为一种关键的中间/输出表征被保留。

### 3.4 占用网络的限制与争议

占用网络并非没有限制。本报告秉持批判性审视，记录以下争议：

第一，**纯视觉占用的可靠性争议**。占用网络从相机生成 3D 占用，其精度受光照、天气、遮挡影响。激光雷达直接测量 3D 距离，在恶劣条件下更鲁棒。Tesla 的纯视觉占用路线在 L4 责任声明上仍受监管质疑——这是第八章将讨论的 AR4 代价之一。

第二，**长尾场景的泛化边界**。占用网络在常见场景表现优异，但对极端长尾场景（罕见障碍物形态、极端天气）的泛化能力，仍是端到端系统可解释性挑战的一部分。

第三，**形态无关性的理论边界**。“占用网络形态无关”是 Tesla 的工程假说，在汽车与人形机器人间已得到一定验证，但能否推广到更广泛形态（四足、轮式、飞行）仍是开放问题——这正是旗舰报告开放问题之一。

### 3.5 占用网络的架构启示

占用网络给出的核心启示是：**跨形态复用的实现，依赖于找到一种“形态无关的中间表示”**。

对于希望借鉴这一思路的玩家，关键在于复制占用网络本身，而在于理解其设计哲学——在感知与决策之间，构建一种不依赖于具体物理形态的、关于物理世界的客观表示（占用场是一种，世界模型是另一种）。一旦拥有这样的表示，上层的决策与控制就可以在不同形态间共享。这是 NVIDIA Cosmos 世界模型、华为 ADS 4.0 WEWA 等架构共同追求的方向，也是物理 AI 走向 AR4 的共同路径。

## 第四章 跨形态复用机理：从汽车到人形机器人

前三章建立了基础——硅（AI5 跨形态共用）、软件（端到端范式可迁移）、感知（占用网络形态无关）。本章将这些线索汇聚为 D3 的核心论点：**Tesla 如何将汽车业务的累积能力跨形态延伸到 Optimus 人形机器人，哪些层级可复用，复用的工程边界与不可复制性源自何处。**这是 AR4”跨形态泛化”判据的实证核心。

理解这一机理的关键起点，是 Tesla 自己在 2025 年 11 月 6 日股东大会上的官方披露——这构成 D3 第四章的论证锚点。

### 4.1 跨形态复用的官方披露：Tesla 11 层共享核心技术

2025 年 11 月 6 日 Tesla 股东大会披露了一份完整的跨形态共享技术清单，正式名称为” Shared core technology”——明确列出从 transport（运输/汽车业务）到 robotics（机器人业务）的 11 层共享核心技术。这份披露是 Tier 1 资料，构成 D3 跨形态复用论证的官方锚点。

**表 4.1: Tesla 官方披露的 11 层共享核心技术（2025.11.6 股东大会）**

编号	共享核心技术	在汽车业务的累积来源	在 Optimus 端的复用形态
1	Actuators（执行器）	Model S/X/3/Y 电机制造	22 DoF/手 + 全身 ~50 执行器，腱驱动
2	Power electronics（功率电子）	Tesla 自研逆变器 + 充电管理	机器人电源管理与执行器驱动
3	Battery（电池）	4680 电芯 + Pack 集成能力	Optimus 2.3 kWh 电池组（沿用 Tesla 电芯制造经验）
4	Manufacturing（制造）	Giga 工厂巨型铸件 + 装配线	Fremont Model S/X 线 2026.5 转产 Optimus，设计产能 1M/年；Giga Texas 设计 10M/年
5	Data communication（数据通信）	车载以太网 + 区域控制器架构	机器人内部高速数据互联
6	Audio system（音频系统）	车内音响系统 + 麦克风阵列	机器人语音交互 + 环境感知
7	Cameras（摄像头）	8 路 Autopilot 摄像头	8 个 Autopilot 摄像头（同型号）
8	A14/A15 chips（AI4/AI5 芯片）	HW4/HW5 自研推理芯片	Optimus Bot Brain 采用 AI4/AI5（首批 AI5 优先给 Optimus）
9	Training cluster（训练集群）	Cortex 2.0（100K H100/H200, 250-500MW）	FSD + Optimus 同源训练
10	Neural simulation（神经仿真）	FSD 视频训练 + 仿真环境	Optimus 操作场景仿真

编号	共享核心技术	在汽车业务的累积来源	在 Optimus 端的复用形态
11	Real-world AI (现实世界 AI)	端到端 + 占用网络 + 世界理解	同源端到端栈, “Photon In, Action Out”

来源: Tesla Annual Shareholder Meeting Presentation, 2025.11.6 (Tier 1)。原始幻灯片副标题” Shared core technology” 明确列出 11 项。

值得注意的是, Tesla 在这份官方披露中将 11 项并列展示, 没有进一步分组或层级划分——这是营销话语的合理选择, 但不够工程性。D3 在尊重 Tesla 官方披露的同时, 认为这 11 项内部存在**根本不同的复用机理**——某些项依赖物理工厂与重资产累积, 某些项依赖算法与数据闭环, 某些项则需要组织层面的整合作支撑。这一区分对追赶者的方法论意义重大, 是下一节展开的核心。

#### 4.2 11 层的三组结构: D3 提炼的三组复用机理

D3 在分析 11 层共享技术的工程实质后, 认为它们应被理解为**三组根本不同的复用机理**, 再加上一项 Tesla 官方未明确披露但实际是 11 层得以成立前提的**组织层**。

表 4.2: D3 对 Tesla 11 层共享技术的三组结构提炼

组别	涵盖 11 层中的项	复用机理	累积时间	不可复制性来源
物理层	1. Actuators 2. Power electronics 3. Battery 4. Manufacturing 5. Data communication 6. Audio system	物理工厂 + 重资产 + 制造 know-how	10-20 年汽车业务累积	资本规模 + 工厂网络 + 供应链积累
智能层	7. Cameras 8. AI chips 9. Training cluster 10. Neural simulation 11. Real-world AI	算法 + 数据闭环 + 同源训练	5-10 年 FSD 工程累积	数据规模 + 自研芯片 + 端到端范式
组织层 (D3 新增)	跨形态团队整合 FSD-Optimus 合并	组织扁平 + 同源 KPI + 同源技术领导	即时决策但需要组织文化匹配	公司治理 + CEO 单点协调 + 反康威定律

这三组的根本区别在于**追赶者复刻它们的难度结构完全不同**——物理层要求资本规模与重资产积累的时间难以压缩; 智能层要求数据闭环与同源训练的算法演进路径较短但仍需 5 年以上; 组织层在原则上可立即调整但在实践中受公司治理与组织文化深度约束。**这三组结构的清晰划分, 是 D3 对 D2 中 21 家追赶者镜像分析的方法论前提**——后续第九章对大众/丰田/华为/小鹏/小米的镜像, 将基于这三组结构做精确判断 (详见第九章 5 家镜像与 11 层可复用矩阵)。

### 4.3 智能层 5 项的工程机理

智能层涵盖 11 层中的第 7-11 项（摄像头、AI 芯片、训练集群、神经仿真、现实世界 AI），是 D3 第一至三章已深度展开的”FSD 栈跨形态可迁移”机理的工程具体化。本节凝练智能层 5 项复用的核心机理。

**第 7 项 摄像头。**Optimus 采用与 Tesla 车辆同款 8 路 Autopilot 摄像头——同型号、同接口、同标定流程。这看似简单的硬件复用，背后是 Tesla 多年累积的摄像头 ISP（图像信号处理）调优、低照度场景优化、雨雪环境鲁棒性——这些 know-how 在 Optimus 端无需重新工程化。

**第 8 项 AI 芯片。**FSD 推理芯片（HW3 → AI4 → AI5）的自研路径，使 Optimus Bot Brain 可直接复用同代 AI 芯片。第一章已论证：**AI5 首批硅片优先给 Optimus + 超算集群而非车端**，这一资源配置揭示了智能层复用的真实方向——不是同一颗芯片跑两个形态，而是同一软件栈跨两种算力等级（车端 AI4/AI4.5 + 机器人 AI5），通过软件抽象层吸收硬件差异。

**第 9 项 训练集群。**Cortex 2.0（Giga Texas, 100K H100/H200, 250-500MW）服务于 FSD + Optimus 的统一训练任务。同源训练集群意味着两个形态的模型在同一套软件栈、数据管线、训练工具下演进——这是 Tesla 数据闭环（C5）与训练基础设施（C6）作为统一控制点的工程具象。

**第 10 项 神经仿真。**Tesla 自研的仿真环境最初为 FSD 设计——海量驾驶场景的合成生成、长尾事件的可控复现——现在直接扩展到 Optimus 的操作场景仿真。仿真栈的同源性，使两个形态共享同一套”合成数据生成 → 模型训练 → 物理部署 → 真实反馈”的闭环。

**第 11 项 现实世界 AI。**这是智能层最深的一层——端到端范式（V12 起）+ 占用网络（形态无关 3D 表示）+ 世界理解（数千万英里驾驶数据训练）。**Optimus 的核心壁垒不在机械结构，而在其继承自 FSD 的感知与世界模型**——这正是 D3 第三章占用网络与第四章前节论证的核心。

**智能层 5 项的整体含义：**智能层复用的工程门槛，是 5-10 年的 FSD 数据闭环与算法演进——这对追赶者而言不是不可达，但需要的不仅是技术能力，更是规模化数据闭环的车队基础（数百万辆车 + 数年累计驾驶数据）。

### 4.4 物理层 6 项的累积逻辑

物理层涵盖 11 层中的第 1-6 项（执行器、功率电子、电池、制造、数据通信、音频），是 D3 v1.1 之前未充分展开的维度——这是 Tesla 11 层官方披露揭示的关键补强。

**第 1 项 执行器。**Optimus Gen 3 的 22 DoF/手 + 全身约 50 个执行器（腱驱动设计），其制造能力直接来自 Tesla 汽车业务 20 年的电机制造累积——从 Model S 的诱导电机到 Model 3/Y 的永磁同步电机，再到 Cybertruck 的高扭矩驱动系统。Tesla 在精密电机制造上的累积，使 Optimus 执行器的成本结构远低于”为机器人专门设计电机”的同行。

**第 2 项 功率电子。**Tesla 自研逆变器、充放电管理 IC、热管理芯片——这些为汽车业务设计的功率电子能力，直接服务于 Optimus 的电源管理与执行器驱动。功率电子的复用，是机器人轻量化与续航能力的关键工程基础。

**第 3 项 电池。**Optimus 单台 2.3 kWh 电池组，相较 Tesla Model 3 75-82 kWh 电池组规模小约 30 倍。但电芯生产（4680）、Pack 集成、热管理、安全证明的核心 know-how 是同一套。**Tesla 在 4680 电芯上的累积，使 Optimus 电池成本远低于”专门为机器人重新研发电池”的同行**——这是 1X Technologies、Figure、Appttronik 等纯机器人创业公司无法获得的结构优势。

**第 4 项 制造。**这是 11 层中最具结构意义的一项。Tesla 在 Giga 工厂巨型铸件（front + rear gigacasting）、装配线自动化、整车制造工艺标准化上的累积，直接成为 Optimus 量产的物理基础。**Fremont Model S/X 生产线于 2026 年 5 月停产并改造为 Optimus 生产线（设计产能 1M/年），Giga Texas 设计长期产能 10M/年**——这一规模在机器人行业前所未有的，其可行性的相当部分来

自汽车业务摊销的工厂经验。马斯克对 Optimus 单台成本 < \$20K 的目标，其可信度的相当部分建立在“用汽车制造经验生产机器人”的成本结构上——而非独立机器人产线的边际成本结构。

**第 5 项 数据通信。** Tesla 车载以太网 + 区域控制器架构 (zonal architecture) 的工程能力，直接服务于 Optimus 内部高速数据互联。这一复用看似平凡，但实际上是机器人作为“分布式感知-决策-执行系统”的工程基础——没有车载级数据通信能力，机器人内部 50+ 执行器与多路传感器的协同将面临严重瓶颈。

**第 6 项 音频系统。** Tesla 车内音响系统、麦克风阵列、噪声抑制算法的累积，复用于 Optimus 的语音交互与环境感知。这一项在 11 层中相对最次要，但作为“消费级产品”的完整体验组件，仍是 Optimus 商业化部署的必要拼图。

**物理层 6 项的整体含义：**物理层复用的工程门槛，是 **10-20 年汽车业务的重资产累积** —— 工厂网络、供应链、制造工艺、电池产能。这一门槛对追赶者而言远高于智能层——智能层的差距可通过激进投资在 5-10 年压缩，但物理层的工厂建设、供应链建立、制造经验积累，难以在不付出同等时间与资本规模的前提下复制。

#### 4.5 组织层：隐性的第 12 层

Tesla 官方 11 层共享技术清单**未明确列出“组织”**——但 D3 认为，组织整合是 11 层共享得以成立的前提，是隐性的“第 12 层”。

**Tesla 组织整合的关键事实：**2025 年 6 月，Optimus 项目领导权由 Milan Kovac 转交 Ashok Elluswamy——后者同时担任 FSD/Autopilot 副总裁。这一人事安排意味着 **FSD 与 Optimus 团队的事实合并**：同一组织、同一技术领导、同一套工程方法论，同时驱动汽车与机器人两个形态。

**组织层为何是隐性第 12 层：**技术复用如果没有组织复用的支撑，会退化为“两个团队各自维护相似代码”的伪复用——这是 11 层共享技术清单背后的隐性假设。Tesla 通过团队合并，从组织层面保证了“同源”不只是技术口号，而是真实的工程协同。具体而言：

- **物理层 6 项的复用**，需要工厂团队、供应链团队、制造工程团队跨“汽车业务”与“机器人业务”协同——若组织上是两个独立事业部，工厂的跨形态摊销不会自动发生
- **智能层 5 项的复用**，需要软件团队、AI 团队、芯片团队跨形态协同——若组织上是两个独立技术栈，端到端范式不会自动迁移到机器人
- **三组复用的协同效应**，进一步要求 CEO 级别的跨形态资源配置决策——AI5 给 Optimus 不给车的决策、Fremont 转产 Optimus 的决策、FSD-Optimus 团队合并的决策，都是 CEO 直接拍板的资源配置

这是 D1“康威定律的逆向应用”的具体形态——**Tesla 为了实现技术上的跨形态复用架构，主动调整组织结构让同一团队同时负责两个形态**。组织层的整合，是 11 层共享技术从“潜在可能”转化为“工程现实”的关键转换器。

**这一组织层的不可复制性：**跨形态的组织整合在原则上可立即调整（任命一个跨形态副总裁、合并两个事业部），但在实践上受**公司治理与组织文化的深度约束**——传统 OEM（如大众、丰田）的事业部界限、利益分配、考核体系都与“跨形态整合”高度不兼容。这是组织层与物理层、智能层不同的不可复制性来源——不是时间或资本，而是**组织文化与公司治理**。

#### 4.6 跨形态复用的真实成本结构与不可复制性

将 11 层共享技术 + 组织层第 12 层综合后，可对“跨形态复用”的真实成本结构与不可复制性，做出更精确的刻画。

**复用的乘数经济学：**12 层复用的核心经济意义，是**研发成本与重资产成本的跨产品摊销**：

- 传统机器人公司必须独立承担芯片、感知算法、训练基础设施、电机制造、电池产能、工厂建设的全部成本，且只能摊销到机器人这一单一产品线
- Tesla 则将这些成本主要摊销在汽车业务（数百万台年产量）上，机器人作为” 边际复用者” 几乎零增量成本地继承了这套能力
- 这一乘数效应正是垂直闭环 AR4 架构的核心回报：**同一套底层能力，服务于多个物理形态，每个形态都摊薄了其他形态的成本**

马斯克的 Optimus 成本目标（年产百万台规模下单台 < \$20K），其可行性的相当部分建立在这种 12 层跨形态摊销的成本结构上——任何独立机器人厂商无法达到这一成本边界。

**复用的真实门槛与不可复制性来源：**

层组	不可复制性主要来源	追赶者压缩可能性
物理层 (6 项)	资本规模 + 工厂网络 + 供应链 + 制造经验, <b>10-20 年累积</b>	极低——重资产累积时间难以压缩, 资本规模差距结构性
智能层 (5 项)	数据闭环 + 自研芯片 + 端到端范式, <b>5-10 年累积</b>	中等——技术可达性高但需大规模车队数据
组织层 (1 项)	公司治理 + 组织文化 + CEO 跨形态决策能力	看似低（一纸任命）实际极高（组织文化深度约束）

**复用的诚实边界：**跨形态复用有天然上限，并非” 一切皆可复用” ——运动控制层（运动学、动力学、双足平衡、接触动力学、~1kHz 高频实时控制、肌腱驱动执行器特性）几乎不可复用（详见第五章）。但这一不可复用边界的存在，恰恰让 12 层可复用部分的稀缺性显化——**Tesla 在可复用层的领先才构成结构性优势。**

**对其他玩家的核心启示：**跨形态复用的真实图景，不是” 算法或软件层的同源迁移” 那么简单——它是物理层（6 项重资产）+ 智能层（5 项算法/数据）+ 组织层（1 项治理/文化）三组协同的整体能力。任何只想” 用软件层模仿 Tesla” 的尝试，都低估了 11 层共享技术背后的真实门槛。Tesla 之所以是” AR4 唯一完整闭环”，正在于这 12 层的完整协同——而非任何单一维度的强大。这是 D3 第八章 AR4 代价边界论证的物理底层基础，也是第九章 5 家镜像精确判断的方法论锚点。



图 4.1 跨形态复用 12 层栈 — Tesla 官方 11 层共享核心技术 (2025.11.6 股东大会披露) + D3 提炼的组织层第 12 层。三组分类：物理层 6 项 (执行器、功率电子、电池、制造、数据通信、音频) + 智能层 5 项 (摄像头、AI 芯片、训练集群、神经仿真、现实世界 AI) + 组织层 1 项 (FSD-Optimus 团队合并)。底部为不可复用层 (运动控制类)。

## 第五章 Optimus 运动学与不可复用层：差异的工程颗粒度

第四章论证了”可复用”，本章论证”不可复用”——这同样重要。AR4 的跨形态复用不是”一切皆可复用”的乌托邦，而是精确区分”可复用的感知与世界理解”与”不可复用的运动学与控制”。理解这条边界，才能避免对跨形态复用的过度浪漫化。

### 5.1 “大脑”可复用，“小脑”不可复用

一个有用的类比：在跨形态复用中，第四章论证的 Tesla 11 层共享核心技术 + 「第 12 层」组织整合，覆盖了”大脑”（感知、世界理解、高层决策）与”物理硬件”（执行器、功率电子、电池、制造）——但”小脑”（底层运动控制、运动学、双足平衡、接触动力学）几乎不可复用。

原因在于物理形态的根本差异。自动驾驶的运动规划是在低维空间中进行的——车辆在道路平面上的运动，本质上是 2D 平面上的轨迹规划（前进、转向、加减速），自由度低、运动学模型相对简单。而人形机器人的运动控制是一个高维问题——涉及全身数十个自由度的协调、重心动态转移、双足力反馈平衡、防摔倒抗扰控制。这是一套完全独立于驾驶逻辑的全新物理控制体系。

### 5.2 Optimus 的运动学规格

理解不可复用层的颗粒度，需要审视 Optimus 的具体运动学规格。这里必须诚实标注：不同公开来源对 Optimus 自由度的表述存在差异，本报告记录这一不确定性而非强行统一。

**手部自由度**（相对一致的数据）：- Gen 3 手部：每只手 22 个自由度（DoF）- 双手执行器总数：约 50 个（每前臂/手约 25 个）- 相较 Gen 2 手部（每手 11 DoF）翻倍，执行器数约 4.5 倍增长 - 驱动方式：肌腱驱动（tendon-driven）+ 空心杯电机，仿生设计

**全身自由度**（来源不一，标注差异）：- 部分来源：40+ 个自由度 - 部分来源：37 个自由度，手部定位精度 0.08mm - 人手作为参照：约 27 个自由度（操作的”黄金标准”）

**数据性质标注**：Optimus 全身自由度的具体数字在不同公开来源（Tesla 官方、第三方分析、媒体报道）间存在差异，且”Gen 3”称号严格而言特指升级的手部，核心机器人平台（躯干、腿、主计算机、电池）仍延续 Gen 2 设计。本报告以”手部 22 DoF/双手约 50 执行器”作为相对确定的数据，全身自由度标注为”约 37-40+，来源不一”。

### 5.3 不可复用层的工程内容

具体而言，以下层级在 FSD → Optimus 迁移中无法复用，必须重新工程化：

第一，**全身运动学与动力学模型**。机器人的逆运动学（给定目标位姿求关节角）、动力学（关节力矩与运动的关系）是机器人特有的，与车辆的二维轨迹规划毫无共通。

第二，**双足平衡与步态控制**。行走涉及重心动态转移、零力矩点（ZMP）控制、防摔倒抗扰——这是车辆完全不存在的问题。

第三，**精细操作（manipulation）控制**。22 自由度手部的抓取、力控、触觉反馈，涉及车辆完全没有的接触动力学（contact dynamics）。

第四，**控制频率与实时性差异**。机器人的全身控制需要远高于车辆控制的频率（机器人平衡控制常需 1kHz 级，车辆控制频率较低），对实时系统提出不同要求。

### 5.4 可复用与不可复用的边界

将第四章 11 层共享核心技术 + 组织层（第 12 层）框架与本章不可复用层综合，可得跨形态复用的精确边界：

跨形态复用的分界线（可复用积累 | 不可复用执行）

✓ 可复用（12 层积累）	✗ 不可复用（运动控制）
物理层 6 项（重资产/制造）	全身运动学 / 动力学模型
智能层 5 项（算法/数据闭环）	双足平衡与步态控制
组织层 1 项（跨形态团队合并）	精细操作接触动力学
	高频实时控制 (~1kHz)
本质：累积型能力 (可跨形态摊销)	本质：形态依赖的物理执行 (无法跨形态摊销)

(12 层逐项明细与复用成本结构见第八章 8.3 完整框图)

这一边界图揭示了一个关键事实——可复用与不可复用的分界，不是“硬件 vs 软件”的传统区分，而是“累积型能力 vs 形态依赖的物理执行”的区分。

物理层 6 项虽然涉及大量硬件（执行器、电池、功率电子、工厂等），但其复用机理是**累积型能力的摊销**——汽车业务 10-20 年累积的工厂网络、制造经验、电芯产能，能跨形态服务于机器人，因为它们的本质是“能复制的工程能力 + 重资产”。相反，不可复用的运动控制层虽然部分涉及软件（运动学模型、力反馈算法），但其本质是与**特定物理形态紧密绑定的高频实时控制系统**——汽车的 2D 平面轨迹规划与人形机器人的全身高维动力学控制，是两套完全独立的物理控制体系，无法通过累积经验跨形态摊销。

这一分界对追赶者具有重要意义：**Tesla 在 12 层共享上的领先来自 10-20 年累积，难以快速复制；但其在不可复用运动控制层上的优势相对较小**——任何具备机器人控制工程能力的团队（包括 1X Technologies、Figure、Appttronik 等纯机器人公司）都能在这一层与 Tesla 竞争甚至领先。Optimus 真正的结构性壁垒，在于 12 层共享而非运动控制本身。

## 5.5 不可复用层的战略含义

不可复用层的存在，对 AR4 架构有三个重要含义：

第一，**跨形态复用有天然上限**。即使是 Tesla 这样的“12 层共享 + 组织整合”垂直闭环极致案例，也无法做到“100% 复用”。运动控制层的重新工程化，是机器人相较汽车的额外投入。任何宣称“汽车能力可无缝变为机器人能力”的叙事，都忽视了这条不可复用边界。Tesla 自己也必须为 Optimus 投入相当规模的运动控制专门团队、肌腱驱动执行器的机械设计、全身动力学控制算法开发——这些工程量与汽车业务无法摊销。

第二，**可复用层的稀缺性源自三种根本不同的累积机制**。这是 D3 v1.0 / v1.1 未充分展开的方法论维度——可复用 12 层并非同质的“可复用能力”，而是源自三种结构上不同的稀缺机制：

- **物理层稀缺源 = 工厂网络 + 重资产 + 制造 know-how，10-20 年汽车业务累积**。Tesla 的 Fremont 工厂、Giga 上海/柏林/德州工厂、4680 电芯产能、自研逆变器供应链——这些资产的稀缺性来自**资本规模 + 时间**双重门槛。任何竞争者即使有无限资金，也难以在 5 年内建成等量的工厂网络与制造工艺成熟度。
- **智能层稀缺源 = 数据闭环 + 自研芯片 + 端到端范式，5-10 年算法演进累积**。FSD 数千万英里驾驶数

据、HW3→HW4→AI5 三代芯片演进、V11→V14 端到端范式 4 次迭代——这些资产的稀缺性来自**数据规模 + 工程演化路径依赖**。竞争者即使有同等技术能力，也需要数年时间积累等量数据闭环。

· **组织层稀缺源 = 公司治理 + 组织文化 + CEO 跨形态决策能力**。Tesla 的扁平组织、跨事业部 KPI 同源、Musk 跨 8 实体协调——这些资产的稀缺性最特殊：**在原则上可立即调整，但在实践上受公司治理与组织文化深度约束**。传统 OEM 的事业部界限、利益分配、考核体系都与”跨形态整合”高度不兼容；这不是时间或资本可弥补的，而是**组织文化变革的难度**。

三种稀缺机制各有不同的”难以速成”机理——资本 / 时间 / 治理——共同构成 Tesla 在 12 层共享上的结构性领先。这一三分稀缺机制刻画，对追赶者的方法论意义重大：**不同稀缺机制对应不同的应对路径**。对物理层稀缺，追赶者的合理路径是聚焦于自身已有的物理累积（如丰田的制造、华为的 ICT 累积）；对智能层稀缺，合理路径是激进的数据闭环建设与算法范式革命（如小鹏的 VLA 2.0）；对组织层稀缺，合理路径是公司治理与组织结构的主动重塑（如华为的”不下场造车，专注 Tier-1”组织选择本身就是治理决策）。

第三，**这条边界正是旗舰报告”开放问题二”的实证**。“占用网络形态无关”假说在汽车与人形机器人间得到验证（智能层的 Real-world AI 可复用），但运动学层的形态依赖（运动控制层不可复用）同样得到验证。**能否将 12 层共享推广到更广泛形态（四足、轮式、飞行），是 D1 旗舰报告留给后续研究的开放问题**。D3 基于 12 层框架，可对这一推广性问题做出更精细的分层判断：

· **物理层 6 项的推广性分层：**

- **大概率可推广：**Battery (4680 电芯)、Power electronics (逆变器/电源管理)、Data communication (车载以太网架构) —— 这些是与具体形态无关的工程能力，理论上可服务于四足、轮式、飞行平台。
- **部分可推广：**Manufacturing 工厂经验 —— 取决于产量规模与制造工艺相似度。四足机器人若达到百万台年产量级，工厂经验可摊销；小规模特种用途则难以复用。Audio system 类似。
- **形态依赖较强：**Actuators —— 双足机器人的高扭矩肌腱执行器、四足机器人的中小型旋转执行器、飞行平台的高 RPM 电机，机械结构差异显著，制造经验复用度递减。

· **智能层 5 项的推广性分层：**

- **理论上可推广：**Real-world AI (端到端 + 占用网络)、Neural simulation、Training cluster、AI chips —— 这些是 D1 论证的”形态无关感知 + 通用世界模型”，原则上可服务于任意形态。
- **形态依赖较强：**Cameras —— 摄像头数量、位置、视角必须根据形态调整。但摄像头本身的硬件 + ISP 调优 know-how 可复用。

· **组织层 1 项的推广性：**跨形态团队合并机制本身可推广到任意新形态。但需要 CEO 级别的资源配置决策与组织文化匹配。

这一分层判断对 D1 旗舰报告”开放问题二”的回答是清晰的：**12 层共享在原则上可推广，但推广深度因层级不同而差异巨大——智能层可推广性最高，组织层取决于公司治理，物理层中部分项形态依赖较强**。这是 D3 对 D1 框架的具体扩展，也是 archi-intelligence 研究系列在 AR4 → AR5 路径上的方法论预设。

## 第六章 训练基础设施：Dojo → Cortex → Dojo 3 的演化

AR4 架构的”训练-仿真-部署统一闭环”判据中，训练基础设施是闭环的引擎。Tesla 训练基础设施的演化历程，充满了战略反复与务实调整，是理解垂直闭环企业”自研 vs 采购”权衡的绝佳案例。

### 6.1 Dojo 的兴衰（2019-2025）

Dojo 是 Tesla 自 2019 年起公开的自研训练超算项目，曾被马斯克定位为“实现 FSD 与 Optimus 商业化的基石”。其核心是 D1 训练芯片（2021，7nm，362 TFLOPS BF16/CFP8），目标是构建一套完全自研的、为视频训练优化的训练基础设施，摆脱对 NVIDIA GPU 的依赖。

然而，2025 年 8 月，马斯克确认关闭 Dojo 项目并解散团队。其公开理由是：“当所有技术路径都收敛到 AI6 时，Dojo 2 已成为”演化死胡同（evolutionary dead end）”。约 20 名工程师离职创立 DensityAI。这一决策的深层逻辑是：**与其维护一套独立的训练芯片架构（Dojo），不如让训练与推理芯片收敛到同一架构（AI5/AI6）。**

马斯克的一句话精准概括了这一收敛：“Dojo 3 可以说以单板上大量 AI6 SoC 的形式存活下来。”换言之，Dojo 作为独立训练芯片架构死亡了，但其使命——自研训练算力——以”用 AI6 推理芯片堆叠成训练集群”的形式延续。

### 6.2 Cortex：务实的混合方案（2024-）

在 Dojo 式微的同时，马斯克自 2024 年 8 月起转而力推 Cortex——建在 Austin/Giga Texas 的巨型 AI 训练超算集群。与 Dojo 的”全自研芯片”路线不同，Cortex 是一个**务实的混合方案**：主要采用 NVIDIA H100/H200 GPU（规模约 6.7 万至 10 万 H100 等价，不同来源数字有差异），用于 FSD 与 Optimus 的视频训练。

Cortex 的战略意义在于：**它承认了在训练侧，自研芯片（Dojo）的演进速度跟不上 NVIDIA 的迭代节奏，因此务实地采购 NVIDIA GPU 以保证训练算力的即时可用性。**这是垂直闭环企业的成熟之处——垂直整合不是教条，在自研不划算的环节（训练 GPU），果断采购。

**数据标注：**Cortex 集群规模在不同公开来源间从”约 6.7 万 H100 等价”到”约 10 万 H100/H200”不等。本报告标注这一区间差异，不强行统一。

### 6.3 Cortex 2.0 与 Dojo 3 重启（2026）

2026 年，训练基础设施进入新阶段：

**Cortex 2.0：**建于 Giga Texas，首期 250MW 于 2026 年 4 月启动，目标 2026 年中达到 500MW 满产。截至 2026 年 5 月，500MW 全部建设尚未完成。Cortex 2.0 被定位为”决定 FSD 改进速度、Optimus 学习新任务速度、Robotaxi 车队扩张信心”的引擎。

**Dojo 3 重启：**2026 年 1 月，Dojo 3 开发重启——但其形态已变。新 Dojo 3 不再是独立训练芯片架构，而是用 Tesla 自研 AI5/AI6 芯片堆叠构建训练集群，目标减少对 NVIDIA 的依赖。这印证了”Dojo 以 AI6 SoC 单板形式存活”的预言。

Tesla 的 AI 芯片路线图现已延伸至 AI5、AI6、AI7 及更远，采用压缩的**九个月设计周期**。配合的产能投入是惊人的：2026 年资本支出预计超过 200 亿美元，与三星签订 165 亿美元的 AI6 芯片制造协议。

## 6.4 训练基础设施演化的架构启示

Tesla 训练基础设施的演化给出三个启示：

第一，**垂直整合不是教条，而是动态权衡**。Dojo（全自研训练芯片）的关闭与 Cortex（采购 NVIDIA）的兴起，证明即使最坚定的垂直闭环企业，也会在自研不划算的环节务实采购。AR4 不要求“一切自研”，而要求“在关键控制点自研”。

第二，**训练与推理芯片的收敛是效率优化**。Dojo 3 以 AI5/AI6 堆叠形式重生，体现了“训练芯片与推理芯片收敛到同一架构”的战略——这降低了维护两套芯片架构的成本，是垂直闭环的效率体现。

第三，**训练基础设施是 AR4 闭环的真实瓶颈**。Cortex 2.0 的算力规模直接决定 FSD 改进速度与 Optimus 学习速度。这解释了为何 Tesla 愿意投入 200 亿美元级资本支出——在 AR4 架构中，训练算力是“数据→能力”转化的真实瓶颈，也是垂直闭环企业必须自己掌控的关键控制点（对应旗舰报告 C6 仿真与算力控制点）。

## 6.5 SpaceX-xAI 合并后的集团级算力重构（2026.2-2026.5）

2026 年 2 月 2 日，SpaceX 完成对 xAI 的收购（“xAI Merger”，合并估值 \$1.25T；详见 SpaceX S-1 招股书 2026.5.20）。这一合并对 D3 训练基础设施分析的影响是结构性的——**原本归属 xAI 的 COLOSSUS（Memphis TN）与 COLOSSUS II（Memphis TN + Southaven MS）训练集群，现在归 SpaceX corporate：**

- COLOSSUS / COLOSSUS II 当前总算力 ~1.0 GW（SpaceX S-1 Q1 2026 数据：AI Segment nameplate compute draw 1 GW）
- COLOSSUS 部署速度 122 天（行业 benchmark：100MW greenfield 数据中心约 2 年）
- COLOSSUS II 部署速度 91 天，更快
- Grok 5 模型当前在 COLOSSUS II 训练
- SpaceX 2026 年 5 月与 Anthropic 签订 Cloud Services Agreement：\$1.25B/月，至 2029 年 5 月，使用 COLOSSUS / COLOSSUS II 部分容量

**这一合并对 D3 论证的方法论意义：**Tesla 的训练基础设施现在被纳入一个更大的集团级算力体系——Tesla Cortex 2.0（Giga Texas，100K H100/H200，250-500MW）服务 FSD + Optimus；SpaceX COLOSSUS / COLOSSUS II（Memphis + Southaven，~1.0 GW）服务 Grok 5 训练 + Anthropic 客户。**这两套基础设施由 Musk 同一人控制，但分属两个公司法律实体**。Terafab 框架（SpaceX + Tesla + Intel）正是为了打通这两套基础设施的硬件供给。

这一架构具有 D1 未充分讨论的方法论含义：**当 AR4 的训练基础设施跨越多个法律实体而通过单一 CEO 协同时，“集团级 AI 算力主权”成为比“单一公司算力主权”更高维的能力**。其他追赶者若没有这种跨实体集团级协同（如华为虽 ICT 实力强但没有跨“汽车+太空”的集团级算力体系），其训练基础设施的可达上限会受结构性约束。这是 D3 对 D1 框架的具体扩展。

## 第七章 组织层面的复用：FSD 与 Optimus 团队合并的架构含义

第四章已提及 FSD 与 Optimus 团队的合并。本章深入这一组织事实的架构含义——因为在 AR4 架构中，组织结构本身就是架构的一部分。

### 7.1 团队合并的事实

2025 年 6 月，Optimus 项目领导权由 Milan Kovac 转交 Ashok Elluswamy。Ashok Elluswamy 同时担任 FSD/Autopilot 副总裁。这一安排使得 FSD 与 Optimus 由同一技术领导、同一核心团队驱动。

如第四章 4.5 节所论证，这一组织整合是 Tesla 11 层共享核心技术得以成立的**隐性第 12 层**——技术层面的“同源”必须通过组织层面的“合并”才能从设计意图转化为工程现实。本章进一步展开这一组织事实的架构含义。

### 7.2 康威定律的逆向应用

软件架构领域有一条著名的“康威定律”：系统架构会反映设计它的组织结构。Tesla 的 FSD-Optimus 团队合并，是康威定律的**逆向应用**——为了实现技术上的跨形态复用架构，Tesla 主动调整组织结构，让同一团队同时负责两个形态。

这一逆向应用的深刻之处在于：**如果 FSD 与 Optimus 由两个独立团队开发，即使技术上可复用，也会因组织边界而退化为“两套相似但不共享的代码”**。技术复用的前提是组织复用。Tesla 通过团队合并，从组织层面保证了“同源”不只是技术口号，而是真实的工程协同。

### 7.3 对“既造车又造机器人”公司的启示

许多公司试图同时进入汽车与机器人领域（小鹏的 IRON、华为的 Kuafu、小米的 CyberOne 等）。Tesla 的组织经验给出一个关键启示：**技术架构的跨形态复用，必须有组织架构的跨形态整合合作支撑**。

如果汽车团队与机器人团队各自为政，即使两者技术栈相似，也无法实现真正的复用——会出现重复造轮子、接口不一致、世界模型不共享等问题。Tesla 的团队合并，是其跨形态复用能够“动真格”的组织保障。这是其他玩家在追求“汽车+机器人”双线时容易忽视的隐性要求。

### 7.4 组织复用的代价

组织复用并非没有代价。FSD 与 Optimus 团队合并意味着：核心人才在两个形态间分配，任一形态的进度都受制于共享团队的带宽；技术领导的注意力被两个形态分散。这是一种“高耦合”组织——收益是深度协同与复用，代价是单点（核心团队/领导）的风险集中。这与垂直闭环架构“全栈耦合”的整体特征一致：高复用伴随高耦合风险。

## 第八章 AR4 的代价与边界：对其他玩家的启示

D3 前七章解剖了 Tesla AR4 垂直闭环的构造。本章作为收尾，回答两个问题：**这套架构的代价是什么？对无法复制 Tesla 的玩家，启示何在？**

### 8.1 AR4 垂直闭环的五重代价

**代价一：极高的资本与人才门槛。**同时垂直整合硅片、OS、应用、云、机器人，需要数百亿美元投入与数千名顶级工程师。Tesla 2026 年承诺资本支出超 250 亿美元（Q1 2026 财报确认，含 Terafab + AI 基础设施 + 六座工厂 + 太阳能制造设备）；与三星签订 165 亿美元 AI6 芯片制造协议；更重要的是 SpaceX S-1 招股书（2026.5.20）披露的硬数据：**SpaceX AI Segment 仅 2026 Q1 一个季度的 capital expenditure 就达 \$7,723M**（约 770 亿元人民币），同期 Space + Connectivity capex 合计仅 \$2,384M——AI 一个季度 capex 是 Space + Connectivity 加起来的 3 倍以上。2025 全年 AI Segment capex \$12,727M。这是只有市值数万亿美元的集团才能支撑的投入规模，且这个规模仍可能不够——Terafab 1 TW/年算力硬件目标隐含的资本需求远超此数。其他 OEM 若想复制 Tesla AR4 路径，资本规模的差距是结构性的、不可在 18 个月内弥合的（参见第九章对小鹏现金 RMB 47.66B 的对照分析）。

**代价二：全栈耦合的脆弱性。**垂直闭环的任何一层失败，可能连锁影响全栈。Dojo 的失败（虽以收敛形式化解）、HW3 存量车队与新软件的张力，都体现了全栈耦合下”难以局部止损”的脆弱性。

**代价三：纯视觉路线的监管争议。**Tesla 坚持纯视觉（无激光雷达）的占用感知路线，在 L4 责任声明上持续受监管质疑。这一架构决策虽是跨形态复用的前提（统一感知栈），但也是其在欧美高度成熟监管市场的最大风险。

**代价四：责任模型的脆弱性。**端到端神经网络（V12 起）的黑盒特性，使得当系统出现难以解释的故障模式时，责任追究与改进都更困难。这是 AR4 端到端范式的内生挑战，也是“旗舰报告”开放问题一”（端到端可证明安全边界）的核心。

**代价五：CEO 个人风险与组织高耦合。**FSD-Optimus 团队合并的高耦合组织，加上 Tesla 战略高度依赖马斯克个人判断，构成了单点风险集中。这是垂直闭环在组织层面的固有脆弱性。

### 8.2 对其他玩家的三类启示

正如旗舰报告所论证，垂直闭环不可在短期复制。但 Tesla 案例对不同类型玩家有差异化启示：

**对资源充足的科技巨头（华为、Google、小米）：**可借鉴”跨形态复用”的思路，但应选择性复用。华为以鸿蒙微内核 + 盘古大模型 + 昇腾算力底座覆盖车-机器人，是”跨设备生态型”对 Tesla 思路的本土化改造。关键是找到自己的”形态无关表示层”（华为的世界模型、Google 的 Gemini Robotics）。

**对传统 OEM（VW、丰田、BMW）：**不应盲目追求全栈垂直整合（CARIAD 的教训），而应识别自己真正可控的控制点。Tesla 的启示不是”你也要造芯片造机器人”，而是”理解哪些是你必须自己掌控的架构控制点，哪些可以平台合作”。

**对机器人专精玩家（Figure、1X、宇树等）：**Tesla 案例揭示了机器人的核心壁垒不在机械结构（小脑），而在世界理解（大脑）。机器人专精玩家若无 Tesla 级别的真实世界数据，需要寻找替代的世界理解来源——这正是 NVIDIA Cosmos 世界模型、各类基础模型对机器人创业公司的价值所在。

### 8.3 跨形态复用的边界：诚实的总结

D3 的核心结论，是对”跨形态复用”的一个诚实刻画——基于第四章 Tesla 官方 11 层共享核心技术 + 组织层第 12 层的完整框架：

=====

跨形态复用的真实图景（基于 Tesla 11 层共享 + 组织第 12 层）

=====

可复用（12 层共享）	不可复用（运动控制层）
-----	
物理层（6 项）：	运动学/动力学模型
Actuators / Power electronics	双足平衡与步态控制
Battery / Manufacturing	精细操作接触动力学
Data communication / Audio	高频实时控制（~1kHz）
	执行器形态特性
智能层（5 项）：	
Cameras / AI chips / Training	
cluster / Neural simulation /	
Real-world AI	
组织层（隐性第 12 层）：	
FSD-Optimus 团队合并 + 同源 KPI	
复用的前提：	
• 物理：重资产工厂网络 + 10-20 年汽车业务累积	
• 智能：数据闭环 + 自研芯片 + 5-10 年算法演进	
• 组织：CEO 跨形态决策 + 反康威定律团队整合	
复用的边界：	
• 12 层可复用，但运动控制层（~6-8 项）不可复用	
• 复用的乘数经济学：汽车摊薄机器人（Tesla）	
vs 独立机器人产线无摊销（Figure/1X）	
• 可复用层的稀缺性：12 层中无一可被追赶者完整复制	

=====

### 8.4 Tesla 作为 AR4 参照系的最终意义

Tesla FSD-Optimus 统一栈作为当前唯一完整的 AR4 案例，其最终意义不在于”成为标杆供人模仿”，而在于提供一个精确的参照系，让整个产业能够校准对”跨形态物理 AI 平台”的理解：

- 它证明了跨形态复用在工程上可行（11 层共享 + 组织整合），同时划定了其边界（运动控制层不可复用）；
- 它展示了 AR4 的代价（资本、耦合、监管、责任、个人风险），让其他玩家能够理性评估自己是否、以及在多大程度上应该追求 AR4；
- 它揭示了 AR4 的核心控制点（自研推理芯片 C2、数据闭环 C5、训练算力 C6、世界理解），为其他玩家识别”自己的可控控制点”提供了坐标。

正如旗舰报告所言：理解 Tesla，不是为了成为 Tesla，而是为了在自己的资源禀赋与组织能力约束下，做出关于架构控制点的清醒判断。这正是 archi-intelligence 研究 Tesla 这一 AR4 案例的最终价值——不是膜拜，而是校准。

---

## 第九章 镜像章：D3 对 D2 其他 OEM 的论证意义

D3 选择 Tesla 作为 AR4 案例研究的方法论意图，从来不是为了让其他 OEM“学习如何成为 Tesla”——旗舰报告（D1）已论证，垂直闭环要求的资本规模、人才密度、组织扁平度是 20 年累积的产物，不可在 18 个月内复制。Tesla 案例的真正价值，是为追赶者提供一个 AR4 参照系——通过 Tesla 工程颗粒度的清晰刻画，让 D2 中其他 OEM 能精确回答：“当我们不可能成为 Tesla 时，我们的真实位置在哪里？我们的合理路径是什么？”

本章选取 D2 中具有代表性的 5 家 OEM——每家代表一种不同的“AR4 企图路径”——做镜像分析。这 5 家覆盖了 D2 评估的全部主要区域与路径形态：欧洲传统转型（大众）、日本渐进路线（丰田）、ICT 跨界进入（华为 HIMA）、中国新势力（小鹏）、消费电子跨界（小米）。

每家镜像分析的最后，都将基于第四章“Tesla 11 层共享技术 + 组织层第 12 层”框架，给出该 OEM 在 12 层框架下的可复用映射判断。9.6 节将给出 5 家 × 12 层的完整可复用矩阵表，将抽象的“AR4 评级”还原为具体可对照的工程颗粒度。

### 9.1 大众 CARIAD：全球总部 vs 中国速度的双线张力

D2 给大众 Snapshot AR2.0、Roadmap AR2.5。其镜像 Tesla 的关键事实：

CARIAD 全球总部承担 SSP (Scalable Systems Platform) 的统一架构定义，目标 2030 年前后落地——这是一个大众希望成为“自己的 FSD 等价物”的全球统一软件平台。然而，CEA (China Electronic Architecture, 中国电子电气架构；VW 集团总部 2024 年首发新闻稿使用“China Electrical Architecture”，自 2025 年起“China Electronic Architecture”成为官方主流表述) 1.0 已于 2025 年末在中国实现 SOP，首发车型 VW ID. UNYX 08 于 2026 年 1 月开始量产，ECU 数量减少约 30%。CEA 由 Volkswagen Group China Technology Company (VCTC)、CARIAD China (1000+ 员工)、小鹏三方共研——这意味着大众实际上接受了“全球 SSP 之外，再做一个仅服务中国市场的并行架构”的双线策略。CARIAD 自研的中国本土 ADAS/AD SoC 已于 2025 年 11 月公布，CARIZON (CARIAD 与地平线 60:40 合资) 承担本土 ADAS 开发，并将 ADAS L2++ 城市 NoA 在 2026 年内推出。

对照 Tesla 的工程颗粒度：大众选择了 Tesla 之外的另一条路径——借力 + 局部自研双线。借力体现为 Xpeng XNGP 授权（2026 年起 5 款车型用 CEA + Xpeng E/E + ADAS stack）；局部自研体现为 CEA 1.0 与 CARIAD 中国自研 SoC。这条路径的方法论意义在于：当全球总部无法在中国速度下完成 Tesla 式的全栈垂直闭环时，“快速借力 + 本地自研”是一个工程上自治的折中。

但这条路径的代价是 D1 已论证的“全栈协同优势缺失”：CEA 与 SSP 是两个并行架构，最终很可能不能融合；Xpeng XNGP 授权的能力存在但 IP 主权不在大众手中；CARIAD 中国自研 SoC 的下一代演进，必须独立承担芯片设计与流片成本——这些成本在 Tesla 集团是被汽车 + Optimus + xAI 共同摊销的，但大众“中国本土自研”的下游用户仅是中国市场。大众的 Roadmap AR2.5 评分，正反映了这种“快速达到 AR2.5 但难以上 AR3”的结构边界——CEA 1.0 解决了“中国市场可用”问题，但未解决“软件 2.0 范式所需的数据闭环与训练基础设施”问题。

12 层框架下的大众：大众在物理层（汽车制造、电池、功率电子、数据通信、音频）有较强累积，但因完全没有人形机器人业务，物理层 6 项在汽车业务内部复用而非跨形态复用——失去了 Tesla “汽车摊薄机器人”的 12 层乘数效应。智能层方面，AI 芯片有 CARIZON 自研路径（局部），训练集群规模显著落后，端到端范式与数据闭环薄弱。组织层方面，CARIAD 总部 vs CARIAD China 的分裂结构，与 Tesla “FSD-Optimus 单一团队”的组织整合相反。综合判断：大众能复制 11 层中的 4-5 项（且都是物理层在汽车业务内部），跨形态乘数效应基本为 0。

## 9.2 丰田 Arene: 渐进式 + fail-operational 与 Tesla 的范式对立

D2 给丰田 Snapshot AR2.0、Roadmap AR2.5。其镜像 Tesla 的关键事实:

**Woven by Toyota 的 Arene 软件平台已于 2025 年 5 月 21 日全球首发，搭载 2026 RAV4**——这是丰田史上首款 SDV (Software-Defined Vehicle)。Arene 由三件套组成: Arene SDK (开发工具包) + Arene Tools (虚拟环境与测试工作流) + Arene Data (数据基础设施)。Lexus ES 在 2026 年内跟进，下一代 EV 全面使用 Arene。Panasonic Automotive Systems 的新 IVI 系统已与 Arene 集成。

**对照 Tesla 的工程颗粒度:** 丰田选择了与 Tesla **范式对立** 的路径。Tesla 是 “**all-in 端到端 + 单点突破**” + **fail-soft** (软失效——L4 之前由人类作为最终安全冗余); 丰田是 “**渐进式 + 多步验证**” + **fail-operational** (容错运行——任何子系统失效后，系统自身仍可降级安全完成任务，无需人类介入)。Arene 的设计理念明确——“它” 将丰田的制造经验应用于现代软件工程”，“分离软件与硬件” 以实现 “零事故未来”。这一理念与 Tesla “**Photon In, Control Out**” 端到端代替 30 万行 C++ 代码的范式革命，处于光谱的两端。

**这条路径的方法论意义在于:** 渐进式 + fail-operational 是 D1 已论证的 “汽车与机器人的失效哲学锚点” ——人命攸关的系统不允许 fail-soft。丰田作为全球年产 1000 万台量级的车企，监管暴露面与品牌责任远高于 Tesla，**它必须选择渐进式**。Arene 的 Automotive Grade Linux 基底 (开源) + Panasonic IVI 集成 (Tier-1 协同) + Woven City 测试场 (受控环境) 的组合，体现了 “在生产规模下保持 fail-operational” 的工程优先级。

**但这条路径的代价是 D1 已论证的 “软件 2.0 范式所需的数据-训练-部署闭环” 的缺失:** 丰田没有 Tesla 数据规模的车队、没有 Cortex 等级的训练算力、没有 FSD V14 等级的端到端神经网络部署经验。Arene 是优秀的 “SDV 基础设施”，但它不是 “AR4 跨形态复用的算法平台” ——丰田的人形机器人项目 (包括 Punyo 等研究) 与 Arene 在算法层未实现深度复用。**Roadmap AR2.5 评分反映了: Arene 解决了 “SDV 起点” 问题，但 AR3 以上需要的 fail-operational + 端到端范式 + 跨形态复用三者协同，丰田当前路径无法在 18 个月内同时达成。**

**12 层框架下的丰田:** 丰田物理层累积全球最强 (年产 1000 万台量级、丰田生产方式 TPS、电池/电机/功率电子全栈)，但因人形机器人项目 (Punyo 等) 与 Arene 在算法层未实现深度复用，物理层 6 项的跨形态摊销空间未被打开。智能层方面，Arene 提供 SDV 基础设施但端到端范式与数据闭环显著薄弱，训练集群规模远落后于 Tesla Cortex 2.0。组织层方面，丰田事业部界限明确 (车 / 机器人 / 工业设备)，与 “康威定律逆向应用” 的跨形态合并组织相反。综合判断: 丰田 12 层框架下，物理层有相当复用潜力但未激活，智能层在 Arene 范畴内有积累但跨形态延展弱，组织层与 Tesla 路径不兼容。

## 9.3 华为 HIMA: AR4 第二路径 (ICT 进入汽车 vs 车企走向 ICT)

D2 给华为 HIMA Snapshot AR4.0、Roadmap AR4+。这是 D2 中**唯一与 Tesla 同等 AR4 评级**的对象，但其镜像 Tesla 的对照价值，恰恰在于 “两条达成 AR4 的截然不同路径”。

关键事实:

- **HIMA 5 个主品牌:** Aito (Seres) / Luxeed (Chery) / Stelato (BAIC) / Maextro (JAC) / Shangjie (SAIC)
- **Jiang 系列 3 个新品牌 (2025-2026 新增):** Aistaland / Yijing (Dongfeng) / Huajing (SGMW) + Qijing (GAC, 2025.9.19) —— **总共合作品牌 8-9 个**
- **累计辅助驾驶里程 8.76 billion km (2026.4)**，接近 10B 关卡 (与 Tesla 同期 10B miles 接近，但里程结构不同)
- **累计装机量 1.4M 车 (2025.12)**，2026 目标 3M + 80+ 车型

- **Qiankun ADS 5.0** 已发布，Aito M9（2026.5.28）首搭，6 颗 LiDAR + 5 mmWave + 12 ultrasonic + 12 cameras
- 新一代 LiDAR（双光路 896 线）—— 全球最高规格量产

**对照 Tesla 的工程颗粒度：**华为是 AR4 路径的**第二条路——ICT 进入汽车**（Tesla 是车企走向 ICT）。两条路径的核心差异：

维度	Tesla 路径	华为路径
起点	整车制造企业 → 走向 ICT	ICT 巨头 → 进入汽车
商业模式	单一品牌全栈垂直闭环	Tier-1 横向授权 8 个车企品牌
数据闭环	单一车队，所有数据归 Tesla	8 个品牌车队，数据归车企但模型归华为
跨形态野心	Optimus 人形机器人，与 FSD 深度复用	不下场做整车与机器人，专注 Tier-1
传感器哲学	纯视觉（去 LiDAR）	视觉 + LiDAR 多融合（896 线 LiDAR）
监管定位	美国 FMVSS + 中国市场补充	中国市场为主，欧洲谨慎试探

**这条路径的方法论意义在于：**华为证明了 AR4 不需要”单一品牌全栈”的组织形态。**横向授权 8 个车企品牌的” Tier-1-Led Alliance” 模式**，能在不下场造车的前提下，实现接近 Tesla 的工程颗粒度——8.76B km 累计里程是这一模式数据规模的实证。

**但这条路径的代价是 D1 已论证的”跨形态复用边界”的缺失：**华为没有 Optimus 等价物，其 AR4 评分基于”汽车 ADS 4.0/5.0 + 鸿蒙座舱 + 三电”的整车级复用，未延展到具身机器人。当 Tesla 把 AR4 的”多体物理 AI 平台”判据演绎到极致（车 + Optimus + xAI 集群三形态共享）时，华为路径在该判据上是有边界的——这并非缺陷，而是商业模式的主动选择（华为明确表态”不下场造车”）。

**对其他 OEM 的方法论启示：**若一家 OEM 既无 Tesla 的 20 年积累，又不具备华为的 ICT 全栈能力，则**最现实的路径是成为华为 HIMA 联盟的成员**（如 SAIC 之于 Shangjie），而非自建 Tesla 等价物。D2 中 BAIC（Stelato）、JAC（Maextro）、Chery（Luxeed）的选择，已经在用脚投票验证这一判断。

**12 层框架下的华为：**华为是 12 层框架最特殊的案例——其复用形态是”提供给授权车企的层” vs”独立保留的层”。提供给车企的层包括智能层 5 项中的 AI 芯片（昇腾）+ 训练集群 + Real-world AI + Neural simulation（4 项强）+ 数据通信（鸿蒙互联）等。独立保留并不下场的层包括物理层的 Manufacturing（不造整车，工厂在合作车企手上）+ Actuators / Battery（不直接做整车机电）。**最关键的缺位是 Tesla 11 层中的 Actuators 项作为”汽车-机器人桥”的角色——华为没有 robotics 业务，物理层 6 项中 5 项的跨形态摊销空间为 0。**组织层方面，华为车 BU + 终端 BG + 智能汽车解决方案 BU 的多 BU 协调代替了 Tesla 单 CEO 跨形态决策。综合判断：华为 12 层中智能层 4-5 项强，物理层因”不下场”主动放弃跨形态摊销，组织层是另一种形态的整合。

#### 9.4 小鹏：最像 Tesla 的中国新势力，晚 5-10 年

D2 给小鹏 Snapshot AR3.0、Roadmap AR4。这是 D2 中**最接近 Tesla 路径**的对象，因此其镜像价值最高。

关键事实（来源：Xpeng SEC Form 6-K + Q1 2026 Earnings, 2026.5.28）：

- **Turing chip 量产**，2026 目标 100 万片，**大众是首个外部商业客户**——这是中国 OEM 自研 AD/ADAS SoC 首次成为西方车企外部采购对象

- VLA 2.0 (Vision-Language-Action 2.0) 2026.3.2 发布, ADAS 里程渗透率 2026.4 突破 50% (Q1 2026 财报)
- VLA “physical Turing test” 2026.3 通过——“乘客几乎分不出是 AI 还是人在开车”
- GX 旗舰 (2026.5.20 SOP): 双 Turing SoC + 全冗余 L4 硬件——中国首款” L4 硬件预装” 量产车
- Robotaxi 试运营 Q3 2026 (广州), 全自动驾驶商业化 2027
- IRON 人形机器人——自研, 2026 年底前量产
- Q1 2026 现金 RMB 47.66B (约 \$6.6B), 2025 Q4 毛利率 21.3%, 净利 RMB 0.38B (首次正净利)

对照 Tesla 的工程颗粒度: 小鹏路径与 Tesla 的对应关系, 呈现高度的”模式同构”:

Tesla 层级	小鹏对应物	时间差距
HW3/HW4/AI4 自研芯片	Turing chip 量产	Tesla HW3 2019 → Xpeng Turing 2026, 约 7 年
FSD V12 端到端	VLA 2.0 端到端 + VLA	Tesla V12 2024.1 → Xpeng VLA 2.0 2026.3, 约 2 年
Optimus 人形机器人	IRON 人形机器人	Tesla Optimus Gen 1 2022 → Xpeng IRON 2026, 约 4 年
Robotaxi (Austin 2025.6 启动)	Robotaxi (广州 Q3 2026 试运营)	约 1.25 年
Cortex 2.0 (100K H100 + 250-500MW)	未披露同等规模训练集群	显著差距

这条路径的方法论意义在于: 小鹏是唯一一家在中国新势力中走”Tesla 完整路径”的厂商——自研 AD/ADAS 芯片、端到端 VLA、L4 硬件预装、Robotaxi、人形机器人。晚于 Tesla 5-10 年, 但路径同构。Tesla 提供的”AR4 参照系”对小鹏的方法论价值最高——它告诉小鹏: - 你的 Turing chip 之后的下一代芯片, 需要面对”硬件代际与软件演进张力” (HW3 unsupervised 撤回事件的等价物); - 你的 VLA 2.0 之后的下一代神经网络, 参数规模将超出 Turing 算力包络, 需要更激进的算力代际跃升; - 你的 IRON 机器人, 需要 FSD-IRON 团队合并的组织重构 (Ashok Elluswamy 等价物); - 你的训练基础设施, 需要从”租用”走向”自建 100K GPU 级集群”。

但小鹏路径的真实风险是 D1 已论证的”组织债务”: 小鹏当前 Q1 现金 RMB 47.66B (约 \$6.6B), 仅为 SpaceX 单季度 AI capex (\$7.7B) 的不到 90%。资本规模的差距, 决定了小鹏可以走 Tesla 路径但走不到 Tesla 的部署深度——除非通过 IPO/股权融资进一步规模化, 或通过大众等西方车企的授权合作 (Turing chip) 实现”反向资本输入”。Roadmap AR4 评分对小鹏意味着: 路径同构 + 时间差距, 2027 年触达 AR4 是可能的, 但维持 AR4 所需的资本规模能否持续, 是开放问题。

12 层框架下的小鹏: 小鹏是 12 层框架下唯一同时下场造车 + IRON 人形机器人 + 自研芯片 + 端到端的中国厂商, 理论上能复制 11 层中的多数项。物理层 6 项: Manufacturing (肇庆 + 武汉工厂) + Battery (采购为主, 自研 800V 平台) + Power electronics (小鹏自研) + Data communication + Audio + Actuators (IRON 自研) —— 6 项全部触及, 但每一项的规模都远小于 Tesla (电池规模差 30 倍以上、工厂产能差 10 倍以上、4680 等价物的累积差 5-10 年)。智能层 5 项: Cameras + AI chips (Turing) + Training cluster (规模未达 Tesla) + Neural simulation + Real-world AI (VLA 2.0) —— 5 项全部触及但 Training cluster 是关键短板。组织层: FSD-Optimus 团队合并的等价物——小鹏 IRON 团队与智驾团队的整合程度未公开披露, 是值得追踪的关键节点。综合判断: 小鹏 12 层中复

用项数最完整（接近 Tesla），但每项规模/深度差距 5-10 年代际。

### 9.5 小米：消费电子供应链效率 + 跨界进入的结构性短板

D2 给小米 Snapshot AR3.0、Roadmap AR3+。其镜像 Tesla 的关键事实（来源：Xiaomi HKEX 年报 + Q1 2026 财报，2026.5.26）：

- 2025 全年交付 411,082 台 (+200.4% YoY) —— 从 0 到 40 万台仅用约 18 个月
- 2025 EV 业务首次年度运营盈利 RMB 0.9B——Q3 2025 首次单季盈利，毛利率 24.3%
- SU7 新代 2026.3.19 发布 → 34 分钟 15000 锁单 / 3 天 30000 锁单
- YU7 GT 在 Nürburgring Nordschleife 创 SUV 圈速纪录 7:22.755
- 2026 交付目标 550K 台
- 累计 R&D 5 年 RMB 105.5B (+37.8%)
- Q1 2026 EV 运营亏损 RMB 3.1B (春节 + SU7 换代影响)

对照 Tesla 的工程颗粒度：小米代表了一条与 Tesla 完全不同的进入路径——消费电子供应链效率 + 品牌跨界 + 速度优先。这条路径在前端市场表现极强（销量、品牌势能、定价权），但在 AR4 关键的工程颗粒度上存在结构性短板：

Tesla 工程要素	小米对应
自研 AD/ADAS SoC (HW3→AI5)	买芯片 (NVIDIA Orin / Thor)
Cortex 2.0 自建 100K GPU 训练集群	租算力 (腾讯云 / 字节跳动云等)
FSD V12-V14 端到端神经网络	小米自研 HAD (Xiaomi Autonomous Driving)，技术追赶中
数据闭环 (数百万车 5 年累计)	2 年累计 60 万车，数据规模显著短板
Optimus 人形机器人 + 跨形态复用	CyberOne 机器人项目，未与汽车深度复用

这条路径的方法论意义在于：小米证明了”消费电子式的供应链效率 + 品牌势能”可以在 2 年内完成”从 0 到 40 万台 + 首次盈利”的市场建立。这是 Tesla 路径之外的另一种合法路径，且其速度甚至超过 Tesla 早期。

但 D3 的镜像意义恰恰在于刻画其结构性短板：当 Tesla 用 AI5 给 Optimus + xAI 集群（而非车）的事实揭示了 AR4 路径下”芯片是集团级 AI 战略下游”时，小米的”买芯片 + 租算力”路径意味着其 AI 算力主权不在自己手中。这一短板在 AR3 以下不显化 (NVIDIA Orin 足以支撑当前 ADAS)，但在 AR3+ 向 AR4 的跃迁中将成为关键瓶颈——追赶 Tesla AR4 的关键路径是”集团级 AI 算力主权”，这是小米当前商业模式无法提供的。小米的 Roadmap AR3+ 评分准确反映了这一边界：速度可以让它达到 AR3+，但 AR4 所需的全栈垂直闭环（含算力主权）超出其商业模式的能力包络。

12 层框架下的小米：小米速度最快但 12 层框架下的复用最依赖外部供应。物理层 6 项：Manufacturing（自建北京工厂，但规模远小于 Giga）+ Battery（CATL/比亚迪供应，非自研）+ Power electronics（部分自研）+ Data communication（澎湃 OS 跨设备互联，强项）+ Audio + Actuators（CyberOne 自研但未与汽车深度复用）—— 6 项中有 3-4 项有不同程度复用，但跨形态摊销空间小。智能层 5 项：Cameras（采购）+ AI chips（买，NVIDIA Orin/Thor）+ Training cluster（租，腾讯云/字节）+ Neural simulation（依赖外部仿真平台）+ Real-world AI（小米 HAD 自研中）—— 关

键的 AI 芯片 + 训练集群两项不属于自有控制。组织层：小米手机 BU + IoT BU + 汽车 BU + 机器人项目的多事业部协调，与 Tesla 单 CEO 跨形态决策不同。综合判断：小米 12 层中的复用更依赖”消费电子供应链效率”而非”自研全栈”，AR3+ 在该模式下可达，但 AR4 的”集团级 AI 算力主权”门槛跨不过。

### 9.6 5 家镜像综合：基于 12 层的可复用矩阵

将 5 家镜像的 12 层框架判断整合，可得 D3 对 D2 中其他 OEM 的核心方法论贡献——Tesla 11 层共享技术 + 组织层第 12 层，是 5 家追赶者在 AR3+ 向 AR4 跃迁中差距的精确刻画。

表 9.1: 5 家追赶者 × 12 层可复用矩阵

层别	项目	大众 CARIAD	丰田 Arene	华为 HIMA	小鹏	小米
物理层	1. Actuators	○	○	○	▲	▲
	2. Power electronics	◆	◆	○	▲	▲
	3. Battery	◆	◆	○	▲	○
	4. Manufacturing	◆	◆	○	▲	▲
	5. Data communication	◆	◆	◆	◆	◆
	6. Audio system	◆	◆	◆	◆	◆
智能层	7. Cameras	◆	◆	◆	◆	◆
	8. AI chips	▲	○	◆	◆	○
	9. Training cluster	▲	▲	◆	▲	○
	10. Neural simulation	▲	▲	◆	▲	▲
	11. Real-world AI	▲	▲	◆	▲	▲
组织层	12. 跨形态合并	○	○	○	?	○
跨形态实现	汽→机器人摊销	○	○	○	▲	▲

矩阵符号：◆ 完全实现（与 Tesla 同级或可比）；▲ 部分实现（有积累但规模/深度不足）；○ 未实现或完全缺位；? 信息不足/未公开披露

#### 矩阵的几个核心观察：

第一，物理层（6 项）的覆盖密度与”是否下场造车”高度相关。大众/丰田作为传统车企，物理层 5-6 项有实质复用基础（除 Actuators 因无机器人业务）。华为不下场造整车，物理层 4-6 项几乎完全缺位。小鹏/小米下场造车但电池/Manufacturing 规模差距大。

第二，智能层（5 项）的覆盖密度与”自研 AI 栈”高度相关。华为唯一在智能层全 5 项 ◆ 完整（昇腾 + 训练集群 + Real-world AI + 仿真 + 摄像头）。小鹏 4 项 ◆/▲（Turing + VLA 2.0 + Camera + 训练集群规模短板）。小米因”买芯片 + 租算力”在 AI chips + Training cluster 两项关键缺位。

第三，组织层（第 12 层）是 5 家中 4 家明确未做到的”反康威定律”整合。只有小鹏可能正在演化（IRON + 智驾团队），但整合程度未公开披露。这是 D3 第七章已论证的”AR4 跨形态复用的组织前

提”——5家中没有一家做到 Tesla “FSD-Optimus 团队合并”的等价物。

第四，**最底行”跨形态实现：汽车→机器人摊销”**是5家中的稀缺资产。只有小鹏（IRON）+小米（CyberOne）下场做机器人；其中小米的机器人未与汽车深度复用。**Tesla 12层共享 + 跨形态摊销的乘数效应，在 D2 的 21 家其他 OEM 中无一完整出现。**

**对追赶者的核心方法论结论：**

第一，**借力 + 局部自研路径**（大众 CEA + Xpeng 授权）在 12 层框架下，能覆盖物理层 5-6 项 + 智能层 1-2 项，AR2.5 可达，AR3 困难——因数据闭环与训练基础设施缺失。

第二，**渐进式 fail-operational 路径**（丰田 Arene）在 12 层框架下，物理层覆盖 6 项但跨形态摊销未激活，智能层 1-2 项部分实现——AR2.5 可达，AR3+ 需补齐端到端范式与跨形态复用能力，时间窗口可能不足。

第三，**ICT 跨界进入路径**（华为 HIMA）在 12 层框架下，智能层 5 项完整 + 组织层另一种形态整合，但物理层 4-6 项主动放弃。AR4 可达（D2 已评级），但 AR4 的”多体物理 AI 平台”判据因无机器人业务不完整满足。

第四，**新势力同构路径**（小鹏）在 12 层框架下，覆盖项数最完整（接近 Tesla 全 12 层），但每项规模/深度差距 5-10 年。AR4 可达性最高但维持 AR4 的资本规模是开放问题。

第五，**消费电子跨界路径**（小米）在 12 层框架下，物理层 3-4 项有不同程度复用，智能层关键 2 项依赖外部供应。AR3+ 可达，AR4 的”集团级 AI 算力主权”门槛跨不过。

**综合而言：**每条路径都不可能成为 Tesla，但每条路径都能在其自身 12 层框架的可达性下抵达某个 AR 等级。理解 Tesla 12 层的工程颗粒度，让追赶者能精确判断”我的真实位置在哪里、我的合理路径是什么”——这正是 D3 在 D1（理论框架）与 D2（22 家全景评估）之间的实证锚点价值。

## 第十章 实证章：D3 对 D1 核心观点的支撑与挑战

D3 选取 Tesla 作为 AR4 案例研究的方法论意图之三，是对 D1（《架构的世纪迁徙》旗舰报告）核心观点的实证检验。本章不是对 Tesla 的赞美录，而是借助 Tesla 这一“唯一完整 AR4 闭环”的工程实体，对 D1 提出的关键判断进行临床检验——哪些得到支撑、哪些受到挑战、D1 框架的可证伪点在哪里。这是 archi-intelligence 研究系列学术严肃性的内在要求。

### 10.1 两层结构论的验证：基础设施收敛 vs 控制语义发散

D1 核心论点之一：汽车 E/E 架构的演进呈现两层结构——底层基础设施（计算、网络、电源、热管理）趋同；上层控制语义与证明义务（功能安全 ASIL、SOTIF、UN-R157）发散。

Tesla 案例对这一论点的实证支撑：

**基础设施收敛的实证：**Tesla AI4 → AI5 → AI6 的代际演进，其底层基础设施（双源代工 TSMC + Samsung、LPDDR5X 内存、PCIe 互联）与行业标准趋同——AI5 的 192GB LPDDR5X 与 NVIDIA Hopper / Blackwell 的内存接口规格类似，Terafab 框架（SpaceX + Tesla + Intel）也明确指向“先进制程 + 先进封装”的行业通用栈。换言之，**Tesla 的硅基础设施层，正在收敛到与全球先进算力基础设施相同的物理标准。**

**控制语义发散的实证：**但 FSD 软件栈与 NVIDIA Drive、华为 Qiankun ADS 5.0、Mobileye SuperVision 在上层控制语义上**完全不收敛**。Tesla FSD 是“纯视觉 + 端到端 + fail-soft（L4 之前由人类作为最终安全冗余）”；华为 ADS 5.0 是“视觉 + LiDAR 多融合 + 渐进式”；Mobileye 是“RSS 形式化安全证明 + 责任分配明确”。这三种控制语义对应三套完全不同的功能安全证明义务、监管对接方式、责任分配模型——没有任何收敛趋势。

**对 D1 的验证结论：**两层结构论得到 Tesla 案例的强支撑。但 D3 也指出 D1 未充分论述的一点：当基础设施层收敛到极致（如 Terafab 1 TW/年的算力硬件愿景），控制语义层的发散会进一步加深——因为底层算力的“过剩”将让上层每个玩家更自由地选择自己的安全哲学、感知哲学、监管哲学。**收敛与发散是耦合演化的，不是独立趋势。**这一观察是 D3 对 D1 的方法论贡献。

### 10.2 “AI 是语言能力，非大脑”在 Tesla 端到端的检验

D1 北极星观点：AI（特别是 LLM 与 VLA 类模型）是语言能力，不是大脑；确定性引擎是骨架，AI 是骨架上的语言层。这一论点是 archi-intelligence 研究系列最重要的判断，也是最容易被 Tesla “all-in 端到端”叙事挑战的论点。

Tesla 案例对这一论点的检验：

**表面挑战：**FSD V12 用单一端到端神经网络替代约 30 万行 C++ 控制代码，看似证明“AI 可以替代确定性引擎”，是对 D1 论点的反例。

**深层支撑：**然而，2026.4 Q1 2026 财报电话会议中，马斯克的关键陈述揭示了相反的事实——“FSD v14.3 软件栈在架构上已足以支持 **unsupervised 部署**，剩下的是 **validation 与监管审批**，不是技术能力问题”。同期 Robotaxi 实际部署 3 城 × 每城 2 辆运营、7 城承诺改口“preparations underway”。这一差距揭示：**在 Tesla 自己的判断里，AR4 落地的真实瓶颈不是神经网络能力，而是验证、安全证明、监管协同**——正是 D1 所说的“确定性引擎是骨架”的具体内容。

更具结构意义的证据来自 SpaceX S-1 招股书（2026.5.20）的 Risk Factors 章节，原文：

“The continued improvement of AI model capabilities has historically depended in part on scaling laws, the empirical observation that model performance improves with increased compute, data, and model size, but **there is uncertainty as to how**

**long these scaling relationships will continue to hold.**”（AI 模型能力的持续提升在历史上部分依赖于 scaling laws——即模型性能随算力、数据、模型规模增加而提升的经验观察——但 **这些缩放关系还能持续多久，存在不确定性。**）

这是 Musk 控制下的公司（SpaceX/xAI 是 Grok 模型的运营方）在 IPO 招股书中——**法律责任最高的文档**——对 scaling laws 的公开质疑。当 Musk 自己的公司在 SEC 文件中对 **scaling laws 的持续性表达不确定时**，D1 “AI 是语言能力非大脑”的判断获得了来自 Tesla/xAI 体系内部的支撑。这不是 archi-intelligence 的理论推论，而是 Musk 公司自己写的 Risk Factor。

**对 D1 的检验结论：**D1 北极星观点不仅未被 Tesla “all-in 端到端”反例，反而被 Tesla 自身的部署节奏与 SpaceX S-1 法律文件双重支撑。**AR4 的真实工程意义，是 AI 语言层（端到端神经网络）与确定性骨架（验证、监管、安全证明）的协同——而非 AI 替代骨架。**

### 10.3 架构债务理论的 Tesla 实证

D1 提出**架构债务**概念：当系统复杂度的增长速度超过工具与方法论的演进速度，组织积累的不可见技术负担。

Tesla 案例对这一概念的实证：

**HW3 unsupervised FSD 撤回事件**（2026.4 Q1 财报确认）是架构债务的典型案列。Tesla 自 2019 年起向 HW3 用户出售（FSD 包售价 \$8,000-\$15,000）“硬件足够支持完全自动驾驶”的承诺；2026 年 4 月，这一承诺被撤回。撤回原因不是 HW3 硬件失败，而是软件复杂度（V14 神经网络参数规模 v.s. V12 增长 4.5 倍）超出了 HW3 算力包络的承诺边界。

**这一事件的方法论意义：**架构债务在 Tesla 这种垂直闭环企业里是**完全自负的**——Tesla 必须自行承担约百万辆 HW3 车队的改装成本（免费 computer + camera retrofit + 城市级 microfactory 网络）。在传统“硬件采购”模式下，Tier-1 与 OEM 分担这种债务；垂直闭环企业则独自承担。

**D3 对 D1 的扩展：**Tesla 案例显示，AR4 路径下的架构债务有一个 D1 未充分讨论的维度——**跨代际承诺的债务**。当软件代际节奏（V12 → V13 → V14，约 18 个月）快于硬件代际节奏（HW3 → HW4，约 4 年；HW4 → HW5，约 4 年）时，“软件向后兼容旧硬件”的承诺成为重大债务。Tesla 的 V14 Lite 项目（为 HW3 维护精简版本）与 microfactory 改装网络，是这一债务的具象化。**这一维度对所有走向 AR4 的 OEM 都将适用——D2 中小鹏 Turing chip 之后的下一代芯片、华为 ADS 5.0 之后的下一代版本，都将面临同等结构的代际承诺债务。**

### 10.4 失效哲学：fail-soft vs fail-operational 的 Tesla 定位

D1 提出**失效哲学**作为汽车/机器人与消费电子/云的“根本分水岭”——消费电子可以 fail-soft（蓝屏重启），汽车必须 fail-operational（任何子系统失效后系统自身可降级安全完成任务）。

Tesla 案例的检验：Tesla 在这一光谱上的定位是开放的。

**官方表述：**Tesla 的 unsupervised FSD 设计目标是 fail-operational——FSD v14.3 的“intervention-free streak counter”功能、Robotaxi 在 Austin/Houston/Dallas 的无安全员运营，都意味着 Tesla 在向 fail-operational 移动。

**实证现实：**但 Tesla 的具体技术选择呈现**部分 fail-soft 特征**。HW3 上的 FSD（V11 至 V12.6）明确是 supervised（要求人类作为最终冗余）；HW4 上的 V14 在 supervised 与 unsupervised 之间存在过渡区间——v14.3 “在架构上足以 unsupervised”但 monitoring 与 takeover 机制仍在保留。**Tesla 的工程哲学似乎是“在监管允许下逐步从 fail-soft 转向 fail-operational”——这与丰田 Arene（明确的 fail-operational 优先级）形成对比。**

**对 D1 的验证结论：**D1 失效哲学论点得到 Tesla 案例的部分支撑，但 Tesla 揭示了 D1 未充分讨论的中间地带——fail-soft 与 fail-operational 不是二分，而是连续光谱，且 Tesla 选择了“渐进式从 soft 走向 operational”。这一选择背后的方法论判断是：**neither full fail-soft (consumer electronics) nor full fail-operational (Toyota), but a calibrated migration enabled by data accumulation and monitoring**——通过数据积累与监控逐步建立 fail-operational 的工程基础。

这一中间地带的存在，对 D2 其他 OEM 的方法论意义是：选择 fail-operational 不需要一步到位，可以通过“AR3 阶段 fail-soft + 数据积累 → AR4 阶段 fail-operational”的迁移路径达成。但前提是数据闭环与监控基础设施已经建立——这恰恰是 D2 多数传统 OEM 的短板。

### 10.5 工具范畴的反向定义：追赶者需要什么样的方法论与工具

本节是 D3 对 D1 的方法论延伸——**从 Tesla 案例反向定义“追赶者向上跃迁所需的方法论与工具范畴”**。

#### 前提的层层递进：

第一前提（D1 已论证）：Tesla AR4 路径要求 20 年累积 + 全栈垂直 + 组织扁平 + 数据闭环，**不可在 18 个月内复制**。

第二前提（D2 已评估）：D2 中欧洲、美国、日韩传统 OEM 的 2027 年 Roadmap 多在 AR2.0-AR3.0，仍承诺向更高 AR 演进。换言之——**追赶者必须冲 AR3+，但又不可能成为 Tesla**。

第三前提（D3 第一至九章已实证）：Tesla AR4 路径的工程颗粒度极其精细——硅代际承诺债务、跨形态复用 12 层栈、训练基础设施 100K GPU 集群、组织 8 实体协同。这些都不能在资源受限的追赶者条件下完整复制；第九章 5 家镜像的 12 层可复用矩阵进一步显示，每家追赶者都在某些关键层级上对 Tesla 工程颗粒度做出妥协。

#### 核心矛盾的定义——

追赶者必须在**没有 Tesla 资源**的前提下，做出**接近 Tesla 质量**的架构决策。

这一矛盾，揭示了一个值得整个产业关注的市场空白。

**市场空白的存在：**当前的工业工具与方法论生态，对**详细设计阶段**已有丰富、成熟的支撑——EDA 工具、PLM 平台、ASIL 安全证明工具链、AUTOSAR 兼容的代码生成器——这些工具帮助架构师把已经决定的架构“画清楚、做对、做精”。然而，工具与方法论生态对**概念探索阶段**——架构师在画第一张电路图、写第一行代码、做第一份成本估算之前，需要做出关键架构决策的那一阶段——仍主要依赖个人与组织级的工程直觉，缺乏系统化的工具与方法论支撑。Tesla 通过 20 年的组织级累积，内化了关于“哪些架构决策对哪些维度有什么影响”的工程直觉；追赶者没有 20 年时间，也没有同等的全栈协同基础。

**这一市场空白的方法论意义：**当追赶者面对“在资源受限下做出接近 Tesla 质量的架构决策”这一核心矛盾时，他们真正缺的不是更好的“画图工具”或“数据看板”——这些工具对详细设计阶段有用，对概念探索阶段不直接解决决策质量问题。追赶者需要的，是一类**承担概念探索阶段决策辅助职责的工具范畴与配套方法论**——其精确特征、实现路径、与既有工具的边界关系，是后续工程实践应当回答的问题，不在本研究的覆盖范围内。

D3 的方法论贡献，是清晰刻画这一市场空白的存在——并将其作为追赶者向 AR3+ 跃迁的关键路径之一提出，供产业界、学术界、工具供应商共同讨论。

### 10.6 D1 的可证伪点：Tesla 路径未来可能的失败模式

archi-intelligence 研究系列的学术严肃性，要求我们不仅论证 D1 框架的支撑证据，也要主动列出**D1 框架在 Tesla 案例上可能失败的具体场景**——即“如果未来发生 X，D1 的判断将被部分证伪”。

### 可证伪点 1: scaling laws 的真实终结

SpaceX S-1 已公开质疑 scaling laws 的持续性。若 2027-2029 年间, scaling laws 在自动驾驶/具身机器人领域被实证终结(即更多数据/算力不再带来比例化的能力提升), 则 Tesla 的”统一栈跨形态复用+数据闭环”竞争优势将显著削弱。**D1 论点会被部分证伪的形式:** D1 “AI 是语言能力非大脑”判断是正确的, 但”语言能力”本身的天花板可能比 D1 预期更低——届时 AR4 闭环的价值会重新评估。

### 可证伪点 2: fail-soft 在 unsupervised 部署中的灾难性失败

若 2026-2028 年间, Tesla unsupervised Robotaxi 在某个城市发生一起或多起**重大伤亡事故**, 导致美国 NHTSA 或欧洲 EU 监管全面回撤 fail-soft 路径的合法性, 则 Tesla 的渐进式 fail-soft → fail-operational 迁移路径将被强制中断。**D1 论点会被部分证伪的形式:** D1 “失效哲学是分水岭”判断仍然成立, 但 D3 第 10.4 节论证的”中间地带可行”将被反证——届时丰田 Arene 的 fail-operational 优先级路径将获得监管层面的方法论胜利。

### 可证伪点 3: 组织债务的临界点

Musk 同时控制 Tesla + SpaceX + xAI + X + Neuralink + Boring + Macrohard + Terafab 8 家实体的董事会多数席位。若其中任何一家出现重大治理危机(监管行动、关键人才流失、产品重大失败), 可能引发组织债务的级联效应。**D1 论点会被部分证伪的形式:** D1 “组织扁平是 AR4 前提”判断仍然成立, 但”单一 CEO 跨实体组织复用”的极限可能比 D1 隐含预期更脆弱——届时 AR4 的”组织协同优势”叙事将被重新审视。

### 可证伪点 4: Terafab 的可行性失败

SpaceX S-1 明确: Terafab 当前是”general framework”, “具体项目仍需单独协商”。若 2027-2028 年间 Terafab 未能产出实际可量产的”1 TW/年”算力硬件, 或 Intel 因自身经营问题退出, 则 Tesla 集团级 AI 算力主权战略将受重大打击。**D1 论点会被部分证伪的形式:** D1 “AR4 需要全栈协同”判断仍然成立, 但”自建芯片制造层”的可行性边界可能比 D1 推论更窄——届时 Tesla 路径对其他 OEM 的”参照系”价值将下降。

### 可证伪点 5: 监管路径的根本分歧

若 2026-2028 年间, 欧盟 AI 法案、UN-R157、美国 FMVSS、中国 GB 7258-2024 等监管框架走向显著分化(如欧洲坚持 fail-operational + 形式化证明、中国接受 fail-soft + 持续监控、美国维持当前监管真空), 则 Tesla 的”单一软件栈+多地区部署”模式将被监管碎片化打破。**D1 论点会被部分证伪的形式:** D1 “AR4 是全球收敛趋势”判断将被反证——届时 AR4 可能不再是单一的”全球阶梯”, 而分化为”美国 AR4”、“中国 AR4”、“欧洲 AR4”三个互不可换的子范畴。

**列出这 5 个可证伪点的方法论意义:** D1 框架必须能在未来 3-5 年的工程现实中接受检验。这 5 个可证伪点是 D1 框架学术严肃性的必要伴随——任何理论若不可被未来证伪, 就不能被未来支撑。后续 D1 版本将持续记录这些可证伪点的演进状况, 并在被实证或反证时公开更新框架。

## 结论：AR4 作为参照系的方法论意义

本报告通过对 Tesla FSD-Optimus 统一栈的深度解剖，完成了 archi-intelligence Research Series 的第三个核心任务——将 D1 的 AR0-AR5 理论框架与 D2 的 22 家 OEM 横向评估，通过一个具体工程实体的临床检验完成闭环。

### Tesla 作为参照系的三重价值

**第一重价值是验证 D1 理论框架的实证锚点。** Tesla 案例对 D1 五个核心判断的检验给出了明确结果：两层结构论（基础设施收敛 vs 控制语义发散）得到强支撑，且 D3 发现了 D1 未充分讨论的耦合演化维度；“AI 是语言能力非大脑”判断被 Tesla 自身部署节奏与 SpaceX S-1 法律文件双重支撑；架构债务理论在 HW3 unsupervised 撤回事件中获得具体形态；失效哲学论点呈现出 D1 未充分讨论的“渐进式 fail-soft → fail-operational”中间地带；组织扁平 + 跨形态复用判据得到 Optimus Gen 3 / Cortex 2.0 / SpaceX-xAI 合并的多重支撑。同时，第十章列出的 5 个可证伪点，是 D1 框架学术严肃性的必要伴随——任何理论若不可被未来证伪，就不能被未来支撑。

**第二重价值是为 D2 的 21 家其他 OEM 提供 AR4 参照系。** 第九章对大众、丰田、华为、小鹏、小米五条路径的深度镜像，揭示了一个共同结论：**每条路径都不可能成为 Tesla，但每条路径都能在其自身约束下抵达某个 AR 等级。** 借力 + 局部自研路径（大众）能解决 AR2.5 但难上 AR3；渐进式 fail-operational 路径（丰田）保持了生产规模安全优先级，但跨形态复用能力缺失；ICT 跨界进入路径（华为）能用 Tier-1 横向授权达成 AR4，但跨形态复用边界存在；新势力同构路径（小鹏）路径相同但资本规模差距决定部署上限；消费电子跨界路径（小米）速度极快但 AR3+ 向 AR4 跃迁所需的算力主权超出商业模式包络。**理解 Tesla 的工程颗粒度，让追赶者能精确判断“我的真实位置在哪里、我的合理路径是什么”**——这是 D3 在 D1 与 D2 之间的实证锚点价值。

**第三重价值是识别追赶者方法论与工具生态的市场空白。** 第十章第 10.5 节通过对追赶者核心矛盾的层层推导，指出当前工业工具生态对详细设计阶段（EDA、PLM、ASIL 安全证明工具链等）已有成熟支撑，但对**概念探索阶段**——架构师在画第一张图、写第一行代码之前需要做出关键架构决策的那一阶段——仍主要依赖个人与组织级的工程直觉。Tesla 通过 20 年累积内化了这一能力；追赶者没有 20 年时间。这一市场空白的清晰刻画，是 D3 对产业界、学术界、工具供应商提出的方法论命题——其具体填补路径，留给后续工程实践与产业讨论。

### AR4 参照系的边界

我们清醒认识到 AR4 参照系的方法论边界：

第一，**Tesla 路径不是 AR4 的唯一路径。** 华为 HIMA 已实证 ICT 跨界 + Tier-1 横向授权的 AR4 路径——虽然其跨形态复用边界与 Tesla 不同，但 D2 评分相同（AR4）。AR4 是能力门槛，不是路径定义。本报告对 Tesla 路径的深度刻画，**不构成对其他 AR4 路径的否定。**

第二，**AR4 不是终点。** D1 框架中还有 AR5（可信通用具身智能体），其工程含义包括跨产业的具身 AI 部署、形式化安全证明、跨地域监管协同等。Tesla 当前仍在 AR4 的爬坡阶段，AR5 的实证案例尚不存在。本报告的方法论 not 适用于 AR4 → AR5 的跃迁分析——那是后续 Working Paper 的任务。

第三，**Tesla 案例处于高速演进期。** 本报告数据截止 2026.5.21。Tesla 的硅路线、FSD 部署节奏、Optimus 量产、Cortex 2.0/COLOSSUS 集群规模、SpaceX-xAI-Tesla 集团级协同，都在快速变化。**任何基于本报告的方法论判断，都应配合后续事件持续更新。**

## 致追赶者

对正在 D2 中各自位置上、希望向更高 AR 等级跃迁的 21 家 OEM 与产业参与者，本报告的最终方法论建议是：

**不要试图复制 Tesla 的 AR4 路径**——20 年累积、全栈垂直、组织扁平、\$25B+ 年度 AI capex (SpaceX S-1 Q1 2026 数据) 是 Tesla 路径的不可分割组件。**应当用 Tesla 作为参照系，精确刻画自己的位置与合理路径。**

具体而言：

- 若你是欧洲传统转型厂（大众、奔驰、宝马、Stellantis、Renault）：你的合理路径可能是”借力 + 局部自研双线”（大众 CEA 模式）或”与中国生态合作”（奔驰、宝马在中国市场的部分尝试），目标 AR2.5-AR3.0。强行追求 Tesla 全栈垂直 AR4 是资本规模与组织文化层面不可行的。
- 若你是日韩渐进路线厂（丰田、本田、现代、起亚）：你的合理路径是 Arene 类的 SDV 基础设施 + fail-operational 优先级 + 渐进式数据闭环建立。目标在 AR2.5-AR3.5，时间窗口拉长至 2030 前后。
- 若你是中国新势力（蔚来、理想、小鹏、小米）：你的合理路径可能是 Tesla 路径的同构但晚 5-10 年（小鹏案例）或消费电子供应链 + 品牌势能（小米案例）。AR4 触达可能在 2027-2028 年，但维持 AR4 所需的资本规模能否持续，是结构性开放问题。
- 若你是中国传统厂（BYD、吉利、长城、SAIC、广汽、东风、长安、Geely、奇瑞）：你的合理路径可能是”与华为 HIMA / Qiankun 联盟成员”（如 Stelato、Maextro、Luxeed、Shangjie、Qijing、Yijing、Aistaland、Huajing），借助 Tier-1 横向授权达成 AR3-AR4 的能力供给。
- 若你是美国传统厂（Ford、GM、Stellantis 北美）：你的合理路径需要直面”既无 Tesla 的全栈能力，又无中国生态的协同”的双重约束。当前最务实的判断可能是接受 AR2-AR3 的能力天花板，将差异化焦点放在品牌、客户关系、特定细分市场上。

**对所有追赶者而言，第十章第 10.5 节指出的”概念探索阶段方法论与工具空白”是 D3 最具结构意义的方法论提示**——在画第一张图、写第一行代码之前，关键架构决策的质量决定了追赶者向上跃迁的可达上限。填补这一空白的具体路径，留给产业界、学术界、工具供应商共同探索，但其方向性意义是清晰的：在 Tesla 资源不可复制的现实下，向 AR3+ 跃迁的关键不在于复制 Tesla 的”画清楚”工具链，而在于建立适合追赶者资源条件下”想清楚”阶段的方法论支撑。

## 致 D1 的可证伪承诺

archi-intelligence Research Series 的学术伦理底线是**承认框架的可证伪性**。第十章第 10.6 节列出的 5 个可证伪点（scaling laws 终结、fail-soft 灾难性失败、组织债务临界、Terafab 不可行、监管路径根本分歧）将在后续 Working Paper 版本中持续记录演进状况，并在被实证或反证时公开更新框架。**这不是 D1 框架的弱点，而是其学术严肃性的必要伴随。**

D3 的工作至此结束。我们期待 2027-2028 年的事实，验证或挑战本报告的判断。Tesla 的 AR4 案例会继续演进，21 家追赶者会陆续给出自己的回答，AR4 → AR5 的工程含义会逐步清晰。archi-intelligence Research Series 将持续以学术独立性为底线、以方法论严谨性为标准、以可证伪性为伦理边界，记录和这场架构的世纪迁徙。

## 附录 A: Tesla AR4 关键事实时间线(Tier 1-2 来源)

本附录汇总正文中散落的 Tier 1-2 带日期事实,作为 Tesla AR4 闭环的快速时序参照。所有条目标注来源层级;数据截止至 2026 年 5 月 21 日。

日期	事件	来源层级
2019	HW3(FSD Computer)量产,~144 TOPS,14nm Samsung——首颗全自研推理芯片	Tier 2
2021、2022	AI Day:公开占用网络、Dojo D1 芯片、端到端方向	Tier 2
2023	HW4 / AI4 搭载于 Model S/X 改款,~500 TOPS,Samsung 7nm	Tier 2
2024.1	FSD V12:约 30 万行 C++ 控制代码被单一端到端网络取代 (“Photon In, Control Out”)	Tier 2
2024.8	马斯克转向力推 Cortex(NVIDIA H100/H200 混合集群)	Tier 2
2024.12	FSD V13:HW4 原生分辨率,训练数据较 V12 约 4.2x	Tier 2
2025 年末	AI4.5 stopgap 自 2026 Model Y 起悄然搭载(因 AI5 延期)	Tier 2
2025.6	Robotaxi 在 Austin 上线(首个无监督商业运营城市)	Tier 1
2025.6	Optimus 项目负责人由 Milan Kovac 转至 Ashok Elluswamy(FSD-Optimus 团队事实合并)	Tier 2
2025.8	Dojo 项目关闭、团队解散;使命以 AI6-SoC 堆叠形式延续	Tier 2
2025.11.6	年度股东大会:官方披露 11 层共享核心技术	Tier 1
2026.1	Dojo 3 开发重启(以 AI5/AI6 堆叠的训练集群形式)	Tier 2
2026.2.2	SpaceX 完成收购 xAI(合并估值 \$1.25T);COLOSSUS / COLOSSUS II 归入 SpaceX	Tier 1
2026.2.17	Cybercab 首台下线(确认使用 AI4,而非 AI5)	Tier 1

日期	事件	来源层级
2026.3	Terafab 框架公布(SpaceX + Tesla; Intel 4 月加入; 约 \$25B Austin)	Tier 1
2026.4	Cortex 2.0 一期 250MW 于 Giga Texas 启用	Tier 1
2026.4.15	AI5 tape-out(GDSII 交付 TSMC); 首批优先供 Optimus + 超算, 而非车端	Tier 1
2026.4.18	Robotaxi 扩展至 Houston、Dallas(共 3 城, 每城约 2 辆)	Tier 1
2026.4.22	Q1 2026 财报电话会: HW3 无监督 FSD 正式放弃; 7 城承诺改口为” preparations underway”	Tier 1
2026.5.17	FSD V14.3.3(固件 2026.14.6.6) 推送 early-access 用户(无干预连续里程计数器)	Tier 2
2026.5.20	SpaceX S-1 招股书提交(AI Segment Q1 2026 capex \$7,723M; Risk Factors 含 scaling law 风险提示)	Tier 1
2026.12(目标)	AI6 tape-out 目标	Tier 2
2027 年中(目标)	AI5 车端高量产	Tier 2

## 附录 B: 关于 archi-intelligence 研究系列

### 使命

archi-intelligence 是一家独立学术研究机构, 致力于将架构智能(Architecture Intelligence, AI<sup>2</sup>) 确立为一种研究范式。其使命是通过开放方法论、透明的数据归属与严格的同行评审, 推进跨产业架构工程实践的标准化与可比性。

### 出版模式

- 所有报告永久免费, 以 CC-BY 4.0 许可发布。
- 每篇报告在 Zenodo 存档并获永久 DOI。
- 全部方法论、原始来源与引用清单 100% 公开。

### v1.0 发布: “三部曲” 套件

- 2026-01: 《架构的世纪迁徙》——旗舰报告(本体论、跨产业基准、AR0-AR5 与 AI<sup>2</sup>-ML 框架)。
- 2026-02: 《2026 全球汽车 E/E 架构成熟度评估》——对 22 家 OEM 在双时间维度(Snapshot + 确认 Roadmap)下的 AR 与 AI<sup>2</sup>-ML 评估。
- 2026-03(本报告): 《多形态物理 AI 平台就绪度: Tesla FSD-Optimus 统一技术栈》——对唯一完整的 AR4 垂直整合闭环案例的深度剖析, 作为闭合 D1 框架与 D2 评估的实证锚点。

### 编辑独立性

本研究的方法论选择、案例评估与结论, 均由 archi-intelligence 研究团队独立做出, 不受任何商业利

益、政治立场或地缘偏好影响。(完整利益冲突披露见 Front Matter。)

#### **承诺**

- 持续出版(至少 5 年,每年 2-3 篇 Working Paper);
- 持续修订(每 6-12 个月一个修订版,所有重大更正公开记录);
- 持续开放(全部方法论与来源公开);
- 持续独立(拒绝来自任何被评估对象的资助)。

#### **联系方式**

- 研究:research@archi-intelligence.org
- 更正:corrections@archi-intelligence.org
- 网站:<https://archi-intelligence.org>

本研究由 archi-intelligence 研究团队编撰。 archi-intelligence Research Series · Working Paper 2026-03(中文版) — 全文完 —