

AI 算力和通信芯片快速发展带来的测试挑战!

泰克科技中国区技术经理 余洋

AI 的热潮席卷全球和各行各业!

以 ChatGPT 为代表的 AGI (通用人工智能) 了, 甚至被称之为第四次工业革命的推动者。比尔·盖茨说, “ChatGPT 像互联网发明一样重要, 将会改变世界。”

AI 芯片和 AI 服务器

ChatGPT 的强大让很多人看到了 AI 所带来的无限可能, 国内外互联网公司纷纷入场, 掀起了建设大模型建设的热潮, 一座座数据中心拔地而起, 高算力显卡被炒到了天价, 仍然一卡难求。

2023 年全球最大的社交网络公司 Meta(原 Facebook) 购买了多达 15 万块 NVIDIA GPU, 和微软相当, 而其它公司如亚马逊、甲骨文、谷歌、腾讯等都只拿到了 5 万块左右。2024 年 1 月 18 日, Meta 的扎克伯格发文称, 计划年底前向英伟达再购买 35 万个 H100 GPU 芯片, 从而使该公司的 GPU 总量达到约 60 万个。英伟达的 H100 售价为 2.5 万至 3 万美元, 35 万个 H100 GPU 的支出在 90 亿美元左右。加上与之相配套的 AI 服务器、交换机、光模块和数据中心基础设施的建设和维护, Meta 2024 年在 AI 上的总支出将达到 940 亿美元以上。

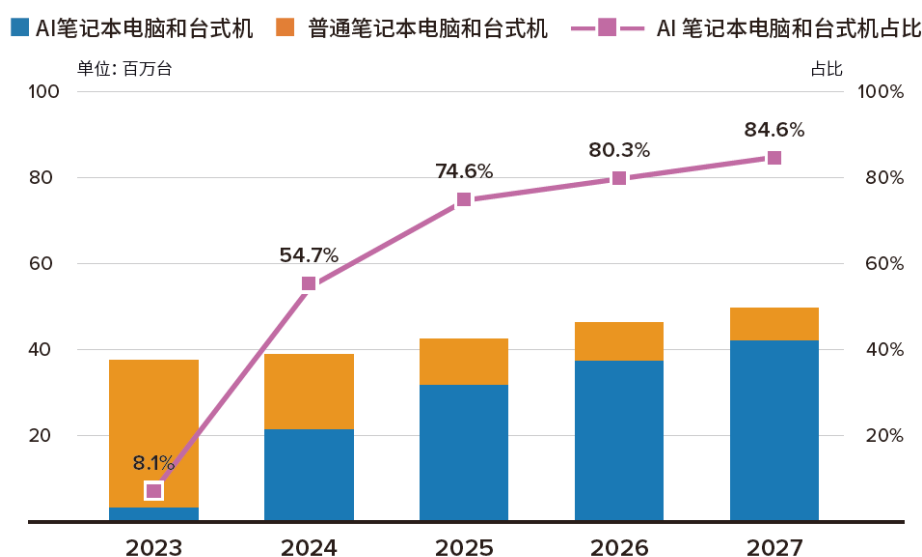
AI PC

2023 年 12 月 7 日, 联想集团与权威国际数据公司 IDC 联合发布了首份《AI PC 产业(中国)白皮书》。PC 作为最强的个人计算平台, 承载着每个人工作和生活中最多的应用场景, 具备实现个人大模型的硬件基础。出于数据安全和隐私保护的考虑, 以及更高效率、更低成本响应用户需求的考虑, 人们既希望获得公共大模型强大的通用服务, 又希望 AI 能够真正理解自己、提供专属个人的服务, 并且能够充分保障个人数据和隐私安全。”未来, 每个个体都可以拥有一个专属于自己的 AI PC, 运行属于自己的“个人大模型”。

IDC 预测, AI PC 在中国 PC 市场中新机的装配比例将在未来几年中快速攀升, 将在 2024 年暴增到 55%, 在 2027 年达到 85%。2024 年将成为 AI PC 元年。英伟达的皮衣老黄也表示十年内 AI PC 将取代传统 PC, 市场规模可达万亿美元级别。

AI PC 市场规模及占比预测

来源: IDC, 2023



注: 本图中的 AI PC 预测数据仅含 AI 笔记本电脑和 AI 台式机, 不含 AI 平板电脑。2023 年的 AI PC 均为 AI Ready 设备。

图 1: AI PC 市场规模及占比预测

AI 手机

无独有偶，AI 大模型的火热，也让手机厂商看到了在软件体验上实现革新的可能。一方面，AI 的进化有望提升智能手机在软件层面上的使用体验，从而带动更多的消费需求。另一方面，跳出硬件互“卷”的怪圈，寻求新的竞争点，打出差异化，也是在新的竞争中获得优势的重要筹码。现在“AI+ 手机”这一概念已经成为了行业共识。1 月 18 日，随着三星新一代旗舰 S24 系列的正式发布，喊出“开启移动 AI 新时代”的口号，在新机中引入视频 AI 处理、AI 聊天机器人、影像画面处理、通话实时翻译等多项 AI 功能，AI 手机正式成为国内外手机厂商共同的“进化趋势”。在此之前，1 月 8 日和 1 月 10 日，国内手机的两场发布会，OPPO 和荣耀，都非常默契地锁定了 AI 大模型这一个焦点。更早之前，小米自演大模型 MiLM，VIVO 的蓝心大模型以及华为的盘古大模型，都已经在 AI 相关的能力落地到了新手机和新系统上。

2 月 18 日早间消息，春节假期后开工首日，OPPO 创始人兼 CEO 陈明永发布一封内部信。他表示，2024 年是 AI 手机元年。未来五年，AI 对手机行业的影响，完全可以比肩当年智能手机替代功能机。AI 手机也将成为继功能机、智能手机之后，手机行业的第三阶段。

同日，魅族科技公布重磅消息：转型 AI 并停止传统手机项目 - 目前还在研发中的魅族机型将全部推倒，All in AI。

AI 赋能汽车

作为继家庭和办公场所之外的“第三空间”，汽车正在变成一个新型智能终端。ChatGPT 到来了之后，车机关系也受到了更多的影响。从整个参与的车企来看，其发展大模型的方向和侧重点并不相同。从功能上来看其主要可以分为以下两类：

一类是用于人工智能交流对话领域，多数应用在智能座舱上。比如，百度的文心一言，目前已经有东风日产、红旗、长城等近十家车企宣布接入。车载大模型语音助手，可以处理完整的对话，如追问，并能保持对前后文的理解，形成较为良好的语音交互体验。驾驶员未来有望通过车载系统完成预订餐厅、预订电影票等任务，极大地丰富智能汽车与人之间的交互体验。

另一类是聚焦智能驾驶的大模型应用。比如毫末智行发布了自动驾驶生成式大模型 DriveGPT，帮助解决认知决策问题，最终实现端到端的自动驾驶；理想汽车自研大模型 MindGPT，摆脱对高清地图的依赖，让汽车做到更接近人类司机的驾驶表现。

这里面既有车企亲自下场，如理想汽车、百度将自家大模型应用于自家汽车产品，也有华为盘古大模型、百度文心一言等用于给其他车企使用。

相较通用大模型来说，垂直领域的大模型训练和使用成本更低，或许会成为更容易实现商业化落地的领域。因此，无论是传统车厂还是新势力，抑或是大模型科技厂商，普遍认为智能汽车最有可能成为率先实现大模型落地的 B 端场景。

AI 无处不在，AI 正在加速爆发

事实上，AI 已经无处不在。在 2024 年 1 月举办的美国消费电子展 CES 2024 上，我们看到了眼花缭乱的 AI 产品。AI 步行鞋帮助人们实现 2.5 倍步行速度，AI 地毯实时收集宠物健康和环境数据，AI 智能腰带帮助盲人进行环境监测和路线导航，AI 枕头帮助用户解决打鼾问题，AI 牙刷检测刷牙习惯和牙齿健康并实时给出语音建议，AI 镜子会告诉你当前的精神状态并提供个性化建议。

你以为这就完了？2024年2月15日，Open AI 在文本模型 ChatGPT、图像模型 Dall-E 大杀四方后，又祭出爆炸性的大杀器 - 视频模型 Sora。Sora 可以根据一段描述性文字生成长达一分钟的视频。无论多天马行空的想法，AI 都可以给你表达出来。Sora 公布的演示视频最令人印象深刻的特点是逼真地模拟物理世界，视频效果和时间长度吊打市面上同类产品如 Runway 和 Pika。2023 年大部分文字生成的 AI 视频长度都在 6 秒以内，业界都把能生成 15 秒视频作为一个里程碑而努力，而这次 Sora 一出手就直接干到了 60 秒，让人们一夜之间又回到了一年多前令人恐惧和焦虑的 GPT-3 时刻。业内分析指出，这将对于广告业、电影预告片、短视频行业带来巨大的颠覆。周鸿祎预测，Sora 的出现，或许意味着 AGI（通用人工智能）的全面落地将从 10 年缩短到 1 年。

AI 算力和通信芯片需要更低的工作电压？

ChatGPT 也好，Sora 也好，大模型训练的背后是由高算力芯片所组成的大规模运算网络，就像开篇提到的，Meta 等巨头一出手就是几十万个高算力芯片，近千亿的投资来建设数据中心。

而作为终端的 AI 产品，比如 AI PC，AI 手机，AI 汽车，AI 智能家居，依赖的是终端产品内的算力芯片。

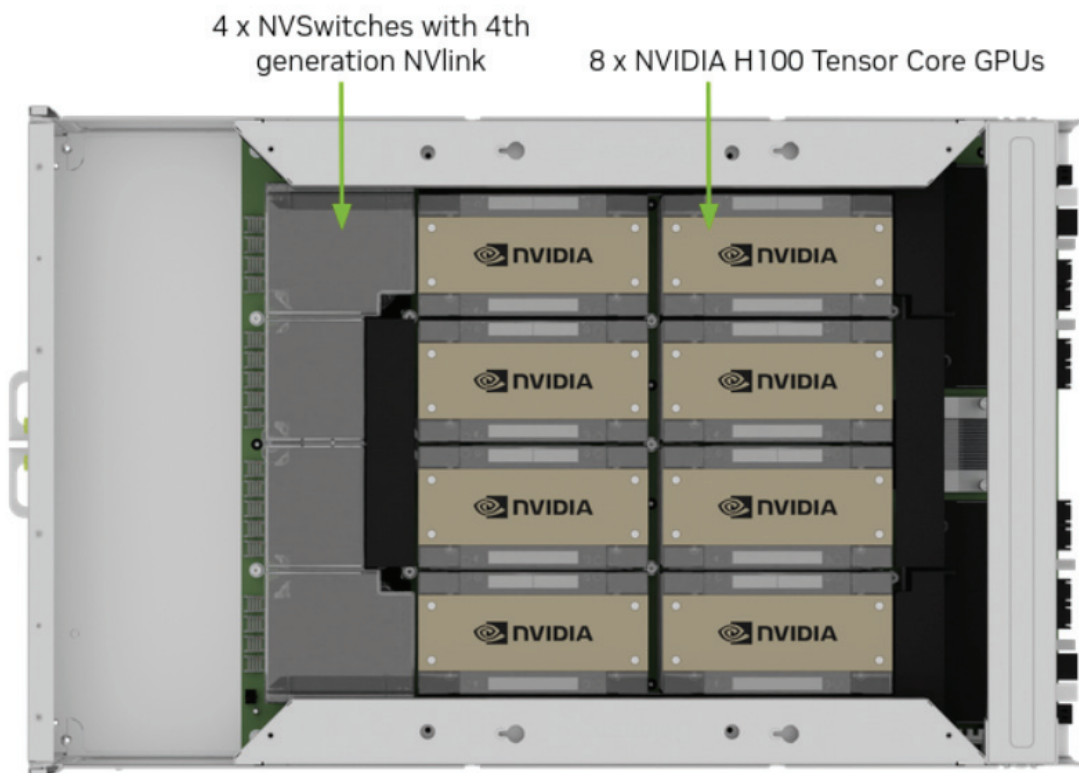


图 2: 典型的 8 x GPU 算力系统

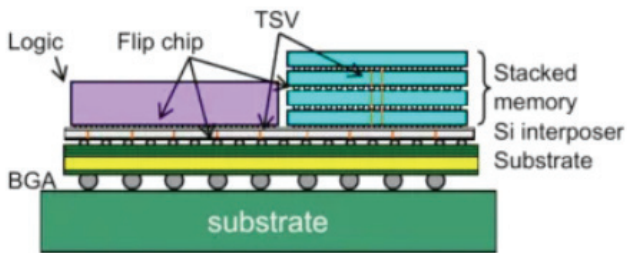


图 4: Chiplet 封装示意图, 存储单元可以多层堆叠而算力单元只能平铺

散热和工艺尺寸一样, 是制约晶体管的密度和规模增加的难题之一。

解决散热的其中一个方案, 就是从源头想办法, 降低电压。使用更低的工作电压, 将每一颗晶体管的功耗降下来, 就可以堆叠更多的晶体管了。

早期的算力芯片工作电压是 5V, 慢慢演化到 3.3V, 1.8V, 1.5V, 到了今天, 算力芯片和高速接口芯片的工作电压基本都在 1V 左右, 甚至更低。这就对电源设计和测量提出了更高的要求。

低电压条件下电源纹波和噪声的测试挑战

电源是算力芯片的能量来源, 是逻辑状态的参考基准。如果电源的纹波和噪声过大, 会给高速变化的逻辑信号上产生大量抖动, 进而产生误码 (注: 误码即错误的码元, 将逻辑 1 当成逻辑 0, 或者将 0 当成 1), 影响芯片的性能, 甚至导致芯片无法正常工作。高速信号验证中非常重要的随机抖动和低频的周期性抖动, 就是由于电源的噪声和纹波所引入的。

电源的纹波和噪声测量, 一直都是电源工程师们最关注的问题之一。算力芯片更低的工作电压, 导致电源留给纹波和噪声的裕度变得更小了, 给设计和测试都带来了难题。

设计上, 算力芯片普遍采用 POL 的降压方式, 将 DC-DC 变压器尽可能靠近负载端, 可以有效避免传输链路上引入的外部干扰。

测试上, 使用更高精度、更低底噪的示波器, 和专用的电源纹波探头, 降低测量系统引入的噪声, 才能更准确地测量电源纹波和噪声。

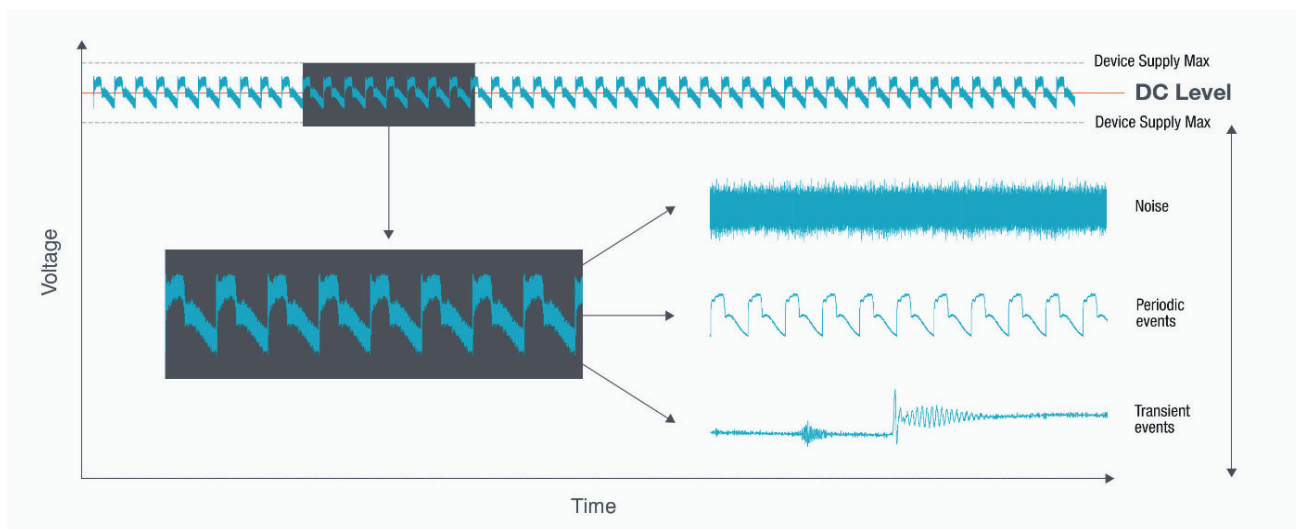


图 5: 电源纹波和噪声

电源的纹波和噪声测量，一直都是电源工程师们最关注的问题之一。算力芯片更低的工作电压，导致电源留给纹波和噪声的裕度变得更小了，给设计和测试都带来了难题。

设计上，算力芯片普遍采用 POL 的降压方式，将 DC-DC 变压器尽可能靠近负载端，可以有效避免传输链路上引入的外部干扰。

测试上，使用更高精度、更低底噪的示波器，和专用的电源纹波探头，降低测量系统引入的噪声，才能更准确地测量电源纹波和噪声。

泰克的解决方案

泰克的 MSO6B 系列示波器的底噪性能非常优异，底噪的有效值在 20MHz 带宽下低至 8.68uV，1G 带宽下低至 51.5uV，是测量电源纹波和噪声很好的选择。



如果电源电压是 1V，示波器的底噪稍微高一点，裕量还有很大空间，是可行的吗？这里需要了解两个问题：

1. 仪器的底噪指标用的都是有效值。而电源纹波和噪声的测量规范，一般都是用峰峰值。峰峰值和测量样本数相关，测量的样本数越多，峰峰值越大，我们可以近似的认为峰峰值是有效值的 10 倍以上。
2. 电源工程师测量底噪和纹波都会使用探头，而探头会引入额外的底噪。

为什么一定要用探头呢？有几个方面的原因，一是探头使用便捷，二是探头提供较高的输入阻抗，对待测电路的影响小，三是探头提供较大的偏置电压，可以在测量噪声和纹波的同时，观察到电源直流电压的变化。尤其当芯片的负载处于动态变化时，电源的直流电压也会随之改变。

示波器加上探头，再去测量一下底噪的峰峰值，你会发现原来底噪还真不小呢。手上有示波器和探头的工程师不妨试试看，将示波器接上探头，不接任何待测信号。在示波器上打开峰峰值测量，测量结果就是系统底噪。

常规的示波器和探头，系统底噪峰峰值在 5 mV 以上。而有些算力芯片和通信芯片，要求电源噪声的峰峰值必须小于 3 mV。测量系统的底噪都这么大，测量结果怎么可能 Pass 呢！

为了更准确的测量电源纹波和噪声，泰克推出了专用的电源轨探头 TPR 系列，20MHz 带宽下的底噪的峰峰值（注意是峰峰值）低至 300 μV ，即便在 4GHz

的全带宽下，底噪的峰峰值依然只有 1.3 mV！而且 TPR 探头还支持高达 60V 的偏置电压，多种多样的探头附件，不仅测得准，用起来还很方便。



指标参数	TPR1000	TPR4000
带宽	1 GHz	4 GHz
偏置电压范围	$\pm 60\text{V}$	
动态范围	$\pm 1\text{V}$	
输入阻抗	50K Ω DC, 50 Ω AC	
耦合方式	DC, LF Reject	
测试精度	1mV	
系统噪声 (与6系示波器一起)	$< 300\mu\text{V}_{\text{Peak-To-Peak}}$ (With 20MHz Bandwidth Limit) $< 1.3\text{mV}_{\text{Peak-To-Peak}}$ (At Full Bandwidth of Scope) Note: With grounded input, set to maximum sensitivity of 1.3mV/Div	
衰减	1.25x	
连接方式	New Browser, Solder-In & Snap-On	

图 6: 泰克 TPR 电源轨探头的核心指标



添加小助手
 申请泰克低噪声测试解决方案
 免费测试机会

更多宝贵资源，敬请登录：WWW.TEK.COM.CN

© 泰克公司版权所有，侵权必究。泰克产品受到已经签发及正在申请的美国专利和外国专利保护。本文中的信息代替所有以前出版的材料中的信息。本文中的技术数据和价格如有变更，恕不另行通告。TEKTRONIX 和 TEK 是泰克公司的注册商标。本文中提到的所有其它商号均为各自公司的服务标志、商标或注册商标。

