

# DIY 深度学习机器（一）：深度学习工作站

## 调研

朱梦

初稿于 2025-07-17，修改于 2025-11-19

### 1. 当前最新 N 卡调研

访问英伟达中文官网<https://www.nvidia.cn>，查看产品 > 显卡，掌握最新架构和最新系列。

表 1 RTX PRO 6000 配置

发布日期	2025-05-22
架构	Blackwell (GB202 满血芯片)
显存	96 GB GDDR7 ECC, 512-bit 总线
算力	FP32: 126 TFLOPS; AI FP4: 4000 TOPS (单卡)
带宽	1797 GB/s (官方 GDDR7 512-bit 理论值)
功耗	600 W (Workstation & Server 版)
汇流排	PCIe 5.0 × 16 (双向 64 GB/s)
外接电源	16-pin PCIe
散热方式	(1) 工作站: 双轴流风扇; (2) 服务器: 涡轮风扇
尺寸	长 267 mm, 高 112 mm, 厚 42 mm
价位 (大陆渠道)	¥74000~¥90000

表 2 RTX 6000 Ada 配置

发布日期	2022-09-20 (官方 GTC 2022 发布)
架构	Ada Lovelace (AD102 专业版)
显存	48 GB GDDR6 ECC, 384-bit 总线
算力	FP32: 91.1 TFLOPS; FP16 Tensor: 1457 TFLOPS; RT Core: 210.6 TFLOPS
带宽	960 GB/s
功耗	300 W
汇流排	PCIe 4.0 × 16 (兼容 PCIe 5.0 插槽, 但速率仍为 PCIe 4.0)
外接电源	16-pin PCIe
散热方式	主动涡轮风扇, 双槽全高设计
尺寸	长 266.7 mm, 高 111.8 mm, 厚 42 mm
价位 (大陆渠道)	¥53000~¥61500

表 3 RTX A6000 配置

发布日期	2020-10-05 (公版首发)
架构	NVIDIA Ampere (GA102-875, 8 nm 定制工艺)
显存	48 GB GDDR6 ECC, 384-bit 总线
算力	FP32: 38.7 TFLOPS
带宽	768 GB/s
功耗	300 W (官方 TDP, 单 8-pin 供电)
汇流排	PCIe 4.0 × 16 (兼容 PCIe 5.0 插槽, 但速率仍为 PCIe 4.0)
散热方式	主动涡轮风扇 (双槽全高, 服务器/工作站通用)
尺寸	长 267 mm, 高 112 mm, 厚 42 mm
价位 (大陆渠道)	¥30000~¥35000

表 4 Tesla H100 配置

发布日期	2022-03-22 (GTC 2022 首发)
架构	Hopper (GH100, TSMC 4N 工艺, 800 亿晶体管)
显存	80 GB HBM3, 5120-bit 总线, 带 ECC
算力	FP64: 30 TFLOPS (SXM) / 24 TFLOPS (PCIe)
带宽	3 TB/s (SXM5) / 2 TB/s (PCIe)
功耗	700 W (SXM5) / 350 W (PCIe)
PCIe	PCIe 4.0 × 16 (兼容 PCIe 5.0 插槽, 但速率仍为 PCIe 4.0)
外接电源	16 -pin PCIe
散热方式	XM5: 冷板液冷 (服务器整机集成)
尺寸	长 267 mm, 高 112 mm, 厚 42 mm
价位 (大陆渠道)	¥190000 (80 GB PCIe 版)

## 2. 适配显卡的 CPU

表 5 适配 RTX 6000 Ada 的 CPU

需求	推荐 CPU	主板芯片组
单机 4 × RTX 6000 Ada, 大模型训练	Xeon W9-3495X	W790 (WS WRX90)
单机 2 × RTX 6000 Ada, 中小模型	Xeon W7-2495X	W790
单机 1 × RTX 6000 Ada, 预算优先	i9-14900K	Z790
边缘/移动场景	N/A	不建议 Core Ultra

表 6 Intel CORE i9-14900KF 与 Intel Xeon W7-2495X 的配置比较

	Intel CORE i9-14900KF 配置	Intel Xeon W7-2495X 配置
核心数	24 核 (8 P-cores + 16 E-cores)	24 核 (Sapphire Rapids-WS)
线程数	32 线程	48 线程
主频	3.2 GHz	2.5 GHz
最大睿频	6.0 GHz (Thermal Velocity Boost)	4.8 GHz (单核) / 3.3 GHz (全核)
基础功耗	125 W	225 W
最大睿频功耗	253 W	270 W
核显	无	无
建议主板	Z790 芯片组	W790 芯片组
接口	LGA 1700, 支持 PCIe 5.0 × 16 + PCIe 4.0 × 4	LGA 4677 (FCLGA4677), 支持 64 条 PCIe 5.0 通道
内存频率	DDR5-5600 / DDR4-3200, 最大支持内存容量为 192 GB	DDR5-4800 MHz (JEDEC), 最大支持内存容量为 2 TB
价位	¥2999	¥18500

### 3. 主板

表 7 技嘉 Z790 主板

型号	技嘉 Z790 AORUS ELITE AX ICE
CPU 接口	Intel LGA1700
PCIe 接口	3 条 PCIe 5.0 × 16
内存接口	DDR5, 最大 7600 MHz, 48 GB × 4
存储接口	1 个 M.2 (CPU 直连) 接口, PCIe 4.0 × 4 / NVMe 协议; 3 个 M.2 (芯片组) 接口, PCIe 4.0 × 4 / NVMe 协议; 6 个 SATA 接口, SATA 3.0 协议
电源接口	24 + 8 pin
板型	ATX (标准型)
主板自身耗电	15~25 W
CPU VRM	400~500 W
价位	¥1999

注：性能按 Z 系列优于 H 系列优于 B 系列排序。

表 8 华擎 W790 主板

型号	华擎 W790 WS 主板
CPU 接口	Intel LGA4677
PCIe 接口	5 条 PCIe 5.0 × 16
内存接口	DDR5, 最大 6800 MHz, 256 GB × 8
存储接口	4 个 M.2 (CPU 直连) 接口, PCIe 5.0 × 4 / NVMe 协议; 1 个 M.2 (芯片组) 接口, PCIe 4.0 × 4 / NVMe 协议; 6 个 SATA 接口, 8 个 SATA 3.0 协议
电源接口	24 + 8 pin
板型	E-ATX (扩展型)
主板自身耗电	25~35 W
CPU VRM	20+2 相 90 A DrMOS (RAA229126 PWM)
价位	¥7084

注：性能按 Z 系列优于 H 系列优于 B 系列排序。

#### 4. 内存条

表 9 不同类型内存条比较

特性	ECC 内存	RECC 内存	台式机内存
纠错功能	纠正单比特错误	纠正单/多比特错误 (更强)	无
稳定性	高	极高	一般
硬件设计	72 位结构, 无寄存器	72 位结构 + 寄存器	64 位标准结构
容量与扩展性	中等 (通常 ≤128GB/通道)	大 (可达 TB 级)	有限 (依赖主板支持)
成本	高	最高	低
主板兼容性	需 ECC 支持的主板 (如服务器主板)	需 RDIMM 接口的主板	通用消费级主板
适用场景	服务器、工作站	高端服务器、数据中心	家用、办公、游戏

表 10 海力士内存条

型号	海力士 ECC
频率	5600 MHz
DDR 代数	第 5 代
容量	48 GB
功耗	4~6 W
价位	¥1299

注：内存条频率的适配需综合考虑 CPU 内存控制器的最大支持频率、主板芯片组的频率上限及超频能力。在非超频情况下，内存频率由两者中较低的限制决定；超频时则需三者（CPU、主板、内存）协同支持。

## 5. 存储硬盘

表 11 三星存储硬盘

维度	三星 PRO 散热片版	三星 870EVO
闪存类型	TCL（市面上很难买到更高级的 SCL 或 MCL）	TCL
接口	M.2	SATA
协议	PCIe 4.0 × 4 / NVMe 协议	SATA 3.0 协议
容量	1TB	1TB
功耗	5.5~7.8 W	2.5~4.0 W
价位	¥879	¥689

注 1：散热片版表示硬盘自带散热片，那么就无需安装主板上原装的“散热马甲”了。

注 2：M.2 硬盘的尺寸、主板 M.2 接口数共同决定了能安装多少个 M.2 硬盘。

## 6. 电源

选购口诀：

- (1) 极致稳定 + 超长质保 → 选海韵/振华/台达。
- (2) 7 × 24 服务器/工作站 → 选台达/TDK-Lambda/施耐德。
- (3) 预算 300~600 元 → 选长城/全汉。
- (4)  $545 \text{ W} \times 1.6 = 850 \text{ W}$  (1.6 倍经验系数)。确认 12VHPWR 原生线或附赠转接线即可。

一句话总结：“家用旗舰看海韵，企业级看台达，预算有限看长城。”

## 7. 机箱

- (1) 全塔机箱：空间更大，通常支持 2 ~ 4 块显卡，适合多卡 SLI/CF 配置，支持 E-ATX 主板和超长显卡。
- (2) 中塔机箱：一般为标准 ATX 主板设计，支持最多 2 ~ 3 块显卡，部分高端中塔也能勉强装下 3 张双槽厚度的显卡，但插槽数量和空间会受限。
- (3) 通常高端显卡搭配全塔机箱，散热更好，方便未来升级和扩展。

表 12 水冷与风冷之间的比较

维度	分体水冷	双塔风冷
散热上限	极高，适合极限超频、高端 CPU/GPU 双冷	高，TDP 280W+，轻松压制中高端 CPU
持续负载	优秀，适合长时间高负载运行	足够，日常/游戏/轻度超频无压力
噪音来源	水泵 + 风扇，调校好可非常安静	风扇为主，高端型号静音优秀
静音上限	更高（可定制低转速水泵 + 风扇）	中等偏上，但风扇多可能略吵
安装难度	高，需要动手能力，走管、注水、排气	中等，体积大但结构简单
机箱要求	需支持大冷排、水箱、水泵空间	需确认 CPU 限高
漏液风险	有，需定期检查水路	无，纯风冷结构
维护需求	高，需换液、清灰、查水泵	极低，偶尔清灰即可
使用寿命	5 年以上（维护得当）	10 年以上，稳定可靠
初期投入	高，动辄千元以上（水泵 + 冷头 + 冷排 + 水箱）	低，200-400 元即可买到旗舰级

表 13 不同型号机箱

机箱型号	类型	显卡限长	支持散热方式	价位
Phanteks 追风者 PK620	全塔	435 mm	风冷 + 分体水冷	¥1399
Phanteks 追风者 ES916	全塔	400 mm	风冷 + 分体水冷	¥6878

## 8. 风扇

风扇吹风方向为：下进上出（热空气质量低于冷空气），前进后出（约定俗成）。因此，上下风扇朝向一致，前后风扇朝向一致。

如何判别风扇哪面为进风，哪面为出风？

- (1) 纸张判别法。运行风扇，纸张远离为出风，吸进则为进风。
- (2) 标签判别法。无标签、无支撑架或框架的一面为进风面，有标签、支撑架或框架的一面为出风面。

下压式 CPU 风扇，朝 CPU 吹风。

塔式散热器，风扇方向和机箱风道统一，自然对流排出热量。