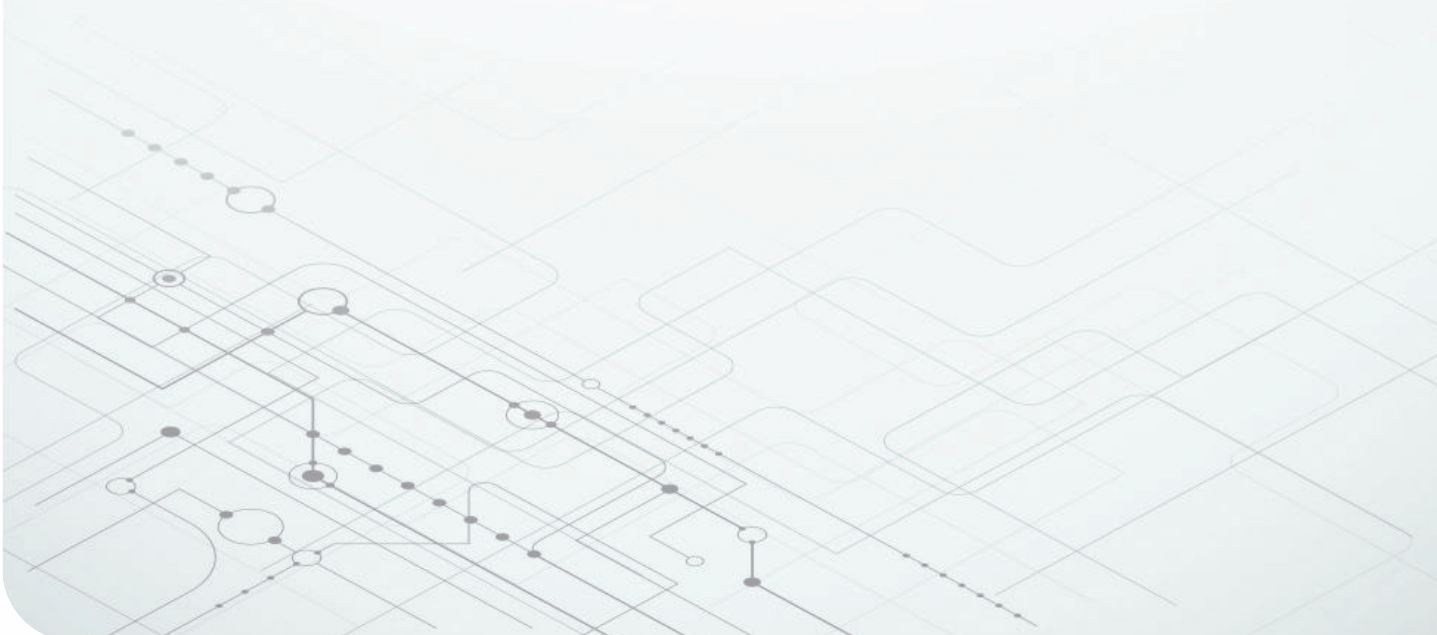


AI 数据中心热管理技术的演变



David Stout
Systems engineer
Building automation - Data center thermal management





内容概览

- **不断上升的热负载：**人工智能 (AI) 数据中心的机架功率密度已从约 5kW 攀升至 100kW 以上，部分 AI 集群预计将达到 1.5MW。
- **冷却架构演进：**随着机架功率超过 100kW，数据中心正从风冷转向液冷和浸没式冷却。
- **预测性热管理：**服务器托盘级液体流量控制与基于边缘 AI 的传感器融合，将帮助数据中心更快检测泄漏并预测故障。

简介

过去几年，数据中心计算设备机架的功率水平持续上升，其驱动力来自大语言模型及其他 AI 工作负载所需的更先进中央处理器 (CPU) 和图形处理器 (GPU) 架构。因此，热管理的重要性也随之提升。

任何**热管理系统**的主要任务在概念上都很简单：从不需要热量的区域吸收热量，并将其转移到可以安全散发热量的位置。随着计算设备机架产生更多热量，要使机架内的处理器保持在可接受的温度范围内运行，就需要采用不同的热管理技术来优化性能。本白皮书概述了热管理系统如何随着功率水平上升而演变，并探讨了为应对不远未来的功率水平可能需要的热管理系统。

相关阅读：“**数据中心产生的热量比以往任何时候都多。为此，散热能力必须同步提升。**”

将机架与家庭中传统的燃木壁炉或燃气壁炉进行比较，是直观了解计算设备机架功率水平及其所产生热量的一种方式。这类壁炉通常在约等于计算设备机架占地面积的范围内，释放出约 10kW 的热量，而这是任何典型空调系统都能轻松处理的负载。但在 50kW 的机架功率水平下，就相当于在同样的占地面积内塞进了 10 个壁炉，而空气冷却已无法在这种热负载下调节室温。一个 1.5MW 的机架相当于在同样空间内放了 150 个壁炉，这确实让人直观地感受到数据中心热管理的挑战。幸运的是，对于相同体积，液冷提供的吸热能力是空气的 3,000 倍。

图 1 展示了冷却挑战升级的速度，将过去、现在和未来的机架功率水平与能够应对它们的架构进行对照，并标出每种架构达到极限的位置。AI 数据中心的机架功率密度已从约 5kW 攀升至 100kW 以上。部分 AI 集群目前每机架已接近 200kW，预计未来几年将高达 1.5MW。这正是冷却架构正从风冷转向液冷以跟上步伐的原因

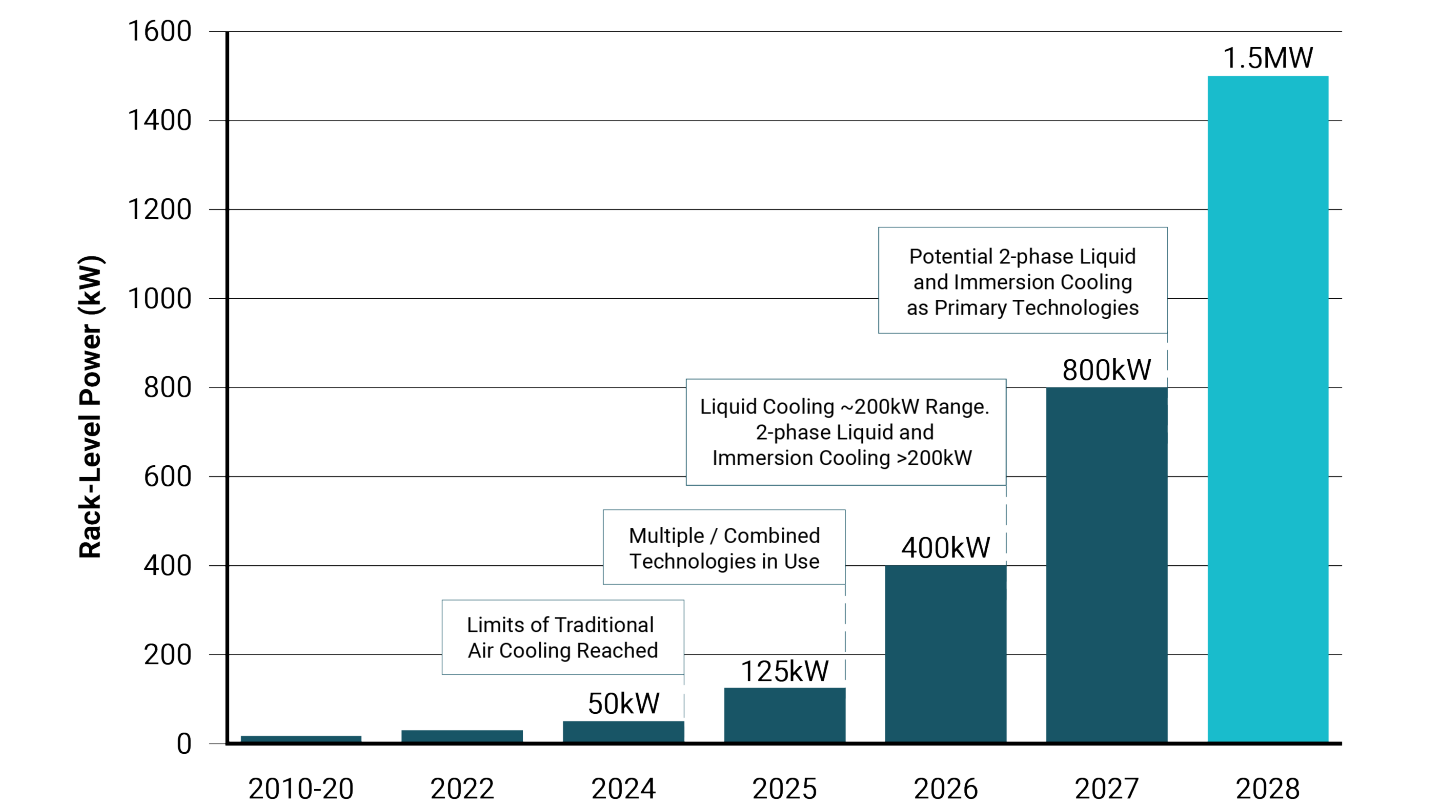


图 1. 机架级功率预测与冷却技术

对比当今的热管理架构

尽管液冷代表着热管理的未来，但混合技术在未来数年内仍将继续服务于该行业。工程师通常将最先进的液冷系统保留给专用于 AI 计算的机架，而其他架构则负责处理标准工作负载。

数据中心主要有三种基于空气的冷却架构，以下按机架功率密度承载能力从低到高排列。每种架构分别描述了用于吸收计算设备机架中所产生热量的介质，以及随后用于将热量传递出数据中心的介质。

空对空

空对空冷却架构的运行方式与住宅中的中央空气系统非常相似，不同之处在于，在数据中心，冷却空气通过计算机架循环，以吸收所产生的热量。然后，热空气通过包含冷制冷剂的线圈，而冷制冷剂会吸收热量并冷却空气以进行再循环。制冷剂将吸收的热量转移到室外，然后释放到环境中，并通过压缩、冷凝和蒸发重新冷却。此过程可以总结为吸热、传热以及散热和冷却的循环。为控制气流，机架通常安装在架高地板上，并以背靠背的方式排列，形成冷通道和热通道。

随着功率水平上升，热通道封闭可防止热空气渗入并抬高由机房空调 (CRAC) 控制的冷通道温度。机房空气处理机组 (CRAH) 负责管理气流，并将热量传递至蒸发器盘管中的制冷剂。热负荷增加后，需要更大的风机风扇和蒸发器线圈，同时

需要更大的服务器风扇气流。在 20kW 至 50kW 的机架功率范围内（假设机房内所有机架配置相似），效率和运行需求要求采用不同的散热方式。

空对液

空对液冷却架构使用与空对空冷却架构相同的方法，但增加了一个后门热交换器 (RDHX)，如图 2 所示。RDHX 的用途是通过计算机架移动的空气中的热量传递到冷液体，同时在空气通过时也对空气进行冷却。RDHX 可以是被动式或主动式，即是否配备额外风扇来辅助气流通过热交换器。

冷却液分配单元 (CDU) 通过泵使吸收了 RDHX 热量的液体循环流动，并将其输送至室外；在室外，冷却塔、干冷器或其他装置将热量释放到环境中，并将液体冷却后重新循环使用。CDU 通常包含一个热交换器，用于将流经机房的液体与流向室外的液体（通常是设施的供水）隔离。这种隔离使得机房内和室外可以使用不同类型的液体。

随着功率水平持续上升，以空气作为吸热介质已成为瓶颈，这促使业界转向使用冷板，从而能够直接从产生热量的部件吸收热量。

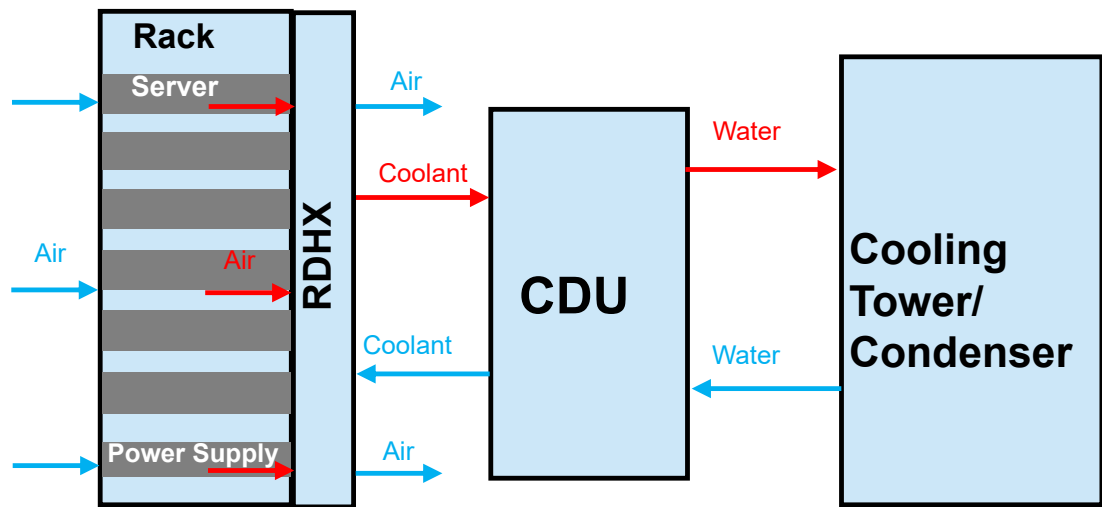


图2. 气液冷却架构

液对空

由于冷板大幅提升了吸热能力，传递热量所需的气流随之减少，使传统 CRAC/CRAH 机组能够继续通过空气传递热量。液对空架构（如图 3 所示）采用散热单元 (HDU)，该单元是 CDU 与主动式 RDHX 的组合。在 HDU 中，CDU 使液体在冷板中循环，冷板安装在机架内各处热源上，以便在尽可能靠近发热点的位置吸收热量。液体流经主动式 RDHX，风扇将机房冷空气吹过盘管，将热量传递给流入热通道的空气。根据功率水平，HDU 可置于机架内、机架排内或作为侧挂单元。

当功率水平攀升至 150kW 至 200kW 范围时，以空气作为传热方式再次成为瓶颈，这促使液对液架构成为当今的主流。

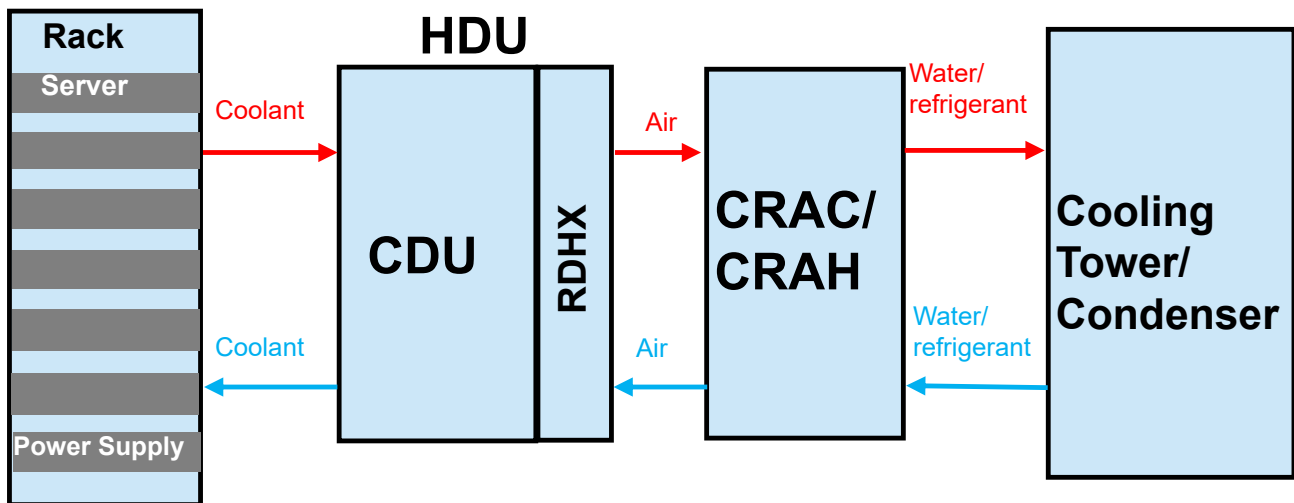


图3. 液对空冷却架构

除了热性能之外，工程师还必须考虑设施要求。空对液架构需要设施水基础设施，并在整个机房空间内铺设大量管路，而液对空架构则可在没有现有水系统的设施中运行，这使其成为改造老旧数据中心时颇具吸引力的选择。空对液、液对空和液对液架构共同构成了当今在用的绝大多数热管理系统。

液体冷却技术的演变

液对液架构（如图4所示）使用直接芯片冷板在热源处吸收热量，并使用冷却液分配单元 (CDU) 使流体循环，将热量从一种液体传递至另一种液体，并从机房空间传递至室外，然后在室外释放。

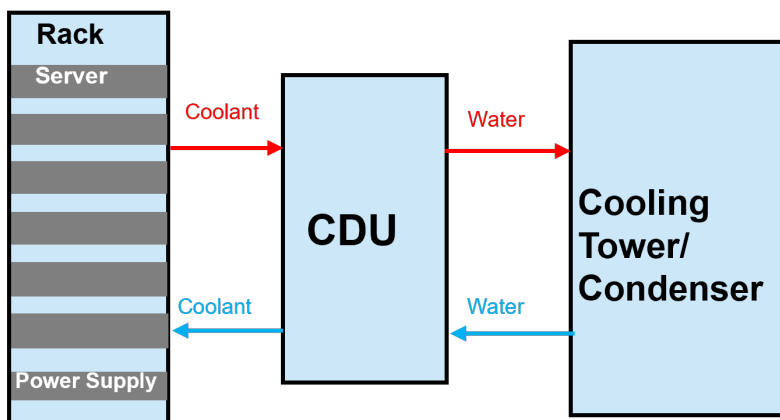


图4. 液对液冷却架构

CDU 是液对液架构中最重要的元件，它保持流体持续流动，从而有效吸收热量，并使各元件保持在预期的温度范围内。CDU 内部的热交换器同样至关重要，它将从冷板吸收热量的低杂质、高比热液体，与室外冷却系统所用的设施水隔离。

泄漏检测的必要性

将液体引入如此靠近电子元件的位置，使泄漏检测变得极其重要。可靠地检测泄漏且误触发少与能否检测到泄漏同样重要，因为服务器正常运行时间取决于此。如今的机架安装采用开闭式阀门，当传感器检测到泄漏时停止液体流动，关闭整个机

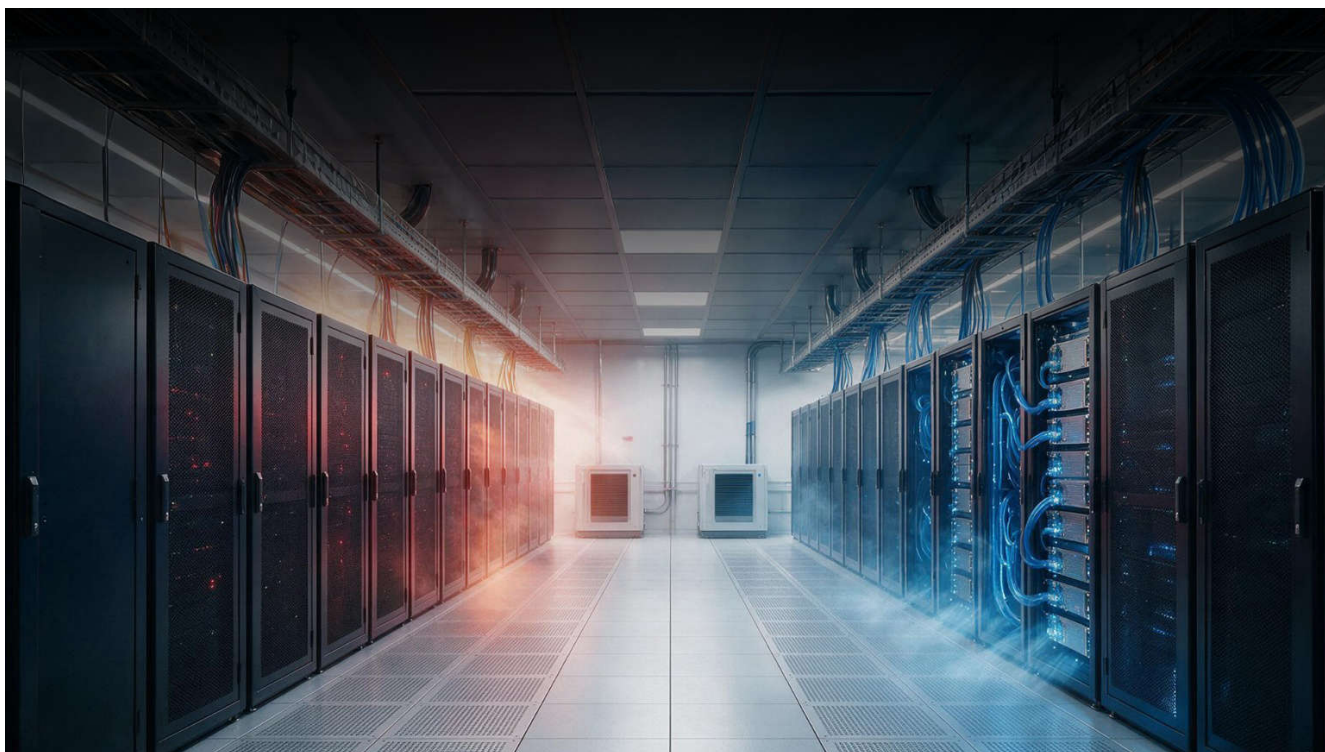
架，直到技术人员找到并修复泄漏源。湿敏电阻绳是最常见的泄漏传感器类型，但它们以精度问题著称，安装时需要小心，避免出现挤压和急弯。工程师正在研究替代方案，包括可检测微泄漏的湿度传感器和电容式泄漏传感器。

其他方法则在 CDU 层面解决泄漏问题。一种方法是重新配置 CDU，使其在负压下将流体抽过系统，而不是推过系统。这样做的效果是吸入空气，而不是迫使流体通过穿孔或松动连接处逸出。另一种方法使用非导电液体，这样泄漏只会造成脏乱，而不会在昂贵设备上造成导电短路。

浸入式冷却

浸入式冷却进一步简化了冷却流程，并满足了专家预测的更高机架功率水平的需求。这种架构将传统机架水平放置，并将其浸没在一槽非导电液体中（不包括金属框架），从而彻底消除因泄漏导致的停机时间。浸入式液冷槽通常集成热交换器和较小的泵来保持液体循环，使该架构具有很高的空间效率，因为它省去了传统机架侧面的歧管和管路。

由于泵和热交换器已内置，CDU 的功能变得自成一体。浸入式冷却的吸热效率大约比直接芯片冷却高出两到三倍，这使位于气候寒冷且设施水温天然较低地区的数据中心，在中等机架功率水平下有可能无需外部冷却塔。



AI 数据中心的未来冷却技术趋势

为高达 1.5MW 功率水平构建的冷却系统，几乎肯定会依赖液对液和浸入式冷却架构。这两种技术都可转换为两相冷却，即利用氟碳介电液（如制冷剂）的低温蒸发，提高吸热和传热效率，从而使这些架构在未来很长一段时间内保持实用价值。

另外两大趋势旨在跟上不断上升的功率密度，并使控制系统响应更迅速：

- 在服务器托盘层面控制液体流量，可更精细地调控 GPU 和 CPU 芯片温度，并在发生泄漏时帮助最大限度保障机架其他部位的服务器正常运行时间。

- 边缘 AI 计算通过传感器融合，将所有系统传感器一起监测和处理，从而提供更快响应时间和更低的延迟，捕捉到单独监测某个传感器或参数时可能遗漏的细微变化。边缘 AI 计算开启了多种可能性，包括提升泄漏检测的可靠性和速度，以及对关键机械元件进行预测性维护。

这些趋势共同作用，使数据中心运营商能够在问题演变为故障之前就清晰洞察到新出现的问题，指向了能够感知状况、预测问题并自动调整的冷却系统。热管理成为数据中心运行中主动参与的一部分，而不是事后才处理的事项。

结语

冷却架构共享许多相同的核心元件。温度传感器在控制温度方面发挥着重要作用。电机控制同样重要，它为这些系统中随处可见的风扇、泵、压缩机和电子阀提供动力。微控制器和微处理器则确保所有部件协同工作。这些只是冷却系统所需半导体器件的几个例子，而 TI 具备良好条件，能够提供覆盖冷却系统各层级的模拟技术和嵌入式处理产品，无论采用何种架构。有关这些冷却架构中所用具体产品领域的详细信息，请参阅“其他资源”部分。

关于作者

David Stout 是一名系统工程师，专注于楼宇自动化系统，包括德州仪器 (TI) 的数据中心热管理。他在为工业半导体市场设计模拟、混合信号及射频电路和系统方面拥有 30 多年的经验。David 拥有路易斯安那州立大学电气工程学士学位，以及德克萨斯大学达拉斯分校电气工程硕士学位。

其他资源

探索与数据中心热管理系统相关的 TI 产品类别：

- 查看[用于逆变器和变频驱动器的 GaN 功率级](#)，以及用于风扇、泵、压缩机和电子阀的[电机驱动器](#)。
- 了解用于泄漏检测和热监测的[温度传感器和其他传感器](#)，以及[精密模数转换器](#)。
- 详细了解保持冷却子系统相互协调且响应迅速的[微控制器和微处理器](#)。

重要声明：本文所提及德州仪器 (TI) 及其子公司的产品和服务均依照 TI 标准销售条款和条件进行销售。建议客户在订购之前获取有关 TI 产品和服务的最新和完整信息。TI 对应用帮助、客户的应用或产品设计、软件性能或侵犯专利不负任何责任。有关任何其它公司产品或服务的发布信息均不构成 TI 因此对其的认可、保证或授权。

所有商标均为其各自所有者的财产。

重要通知和免责声明

TI“按原样”提供技术和可靠性数据（包括数据表）、设计资源（包括参考设计）、应用或其他设计建议、网络工具、安全信息和其他资源，不保证没有瑕疵且不做任何明示或暗示的担保，包括但不限于对适销性、与某特定用途的适用性或不侵犯任何第三方知识产权的暗示担保。

这些资源可供使用 TI 产品进行设计的熟练开发人员使用。您将自行承担以下全部责任：(1) 针对您的应用选择合适的 TI 产品，(2) 设计、验证并测试您的应用，(3) 确保您的应用满足相应标准以及任何其他安全、安保法规或其他要求。

这些资源如有变更，恕不另行通知。TI 授权您仅可将这些资源用于研发本资源所述的 TI 产品的相关应用。严禁以其他方式对这些资源进行复制或展示。您无权使用任何其他 TI 知识产权或任何第三方知识产权。对于因您对这些资源的使用而对 TI 及其代表造成的任何索赔、损害、成本、损失和债务，您将全额赔偿，TI 对此概不负责。

TI 提供的产品受 [TI 销售条款](#)、[TI 通用质量指南](#) 或 [ti.com](#) 上其他适用条款或 TI 产品随附的其他适用条款的约束。TI 提供这些资源并不会扩展或以其他方式更改 TI 针对 TI 产品发布的适用的担保或担保免责声明。除非德州仪器 (TI) 明确将某产品指定为定制产品或客户特定产品，否则其产品均为按确定价格收入目录的标准通用器件。

TI 反对并拒绝您可能提出的任何其他或不同的条款。

版权所有 © 2026，德州仪器 (TI) 公司

最后更新日期：2025 年 10 月