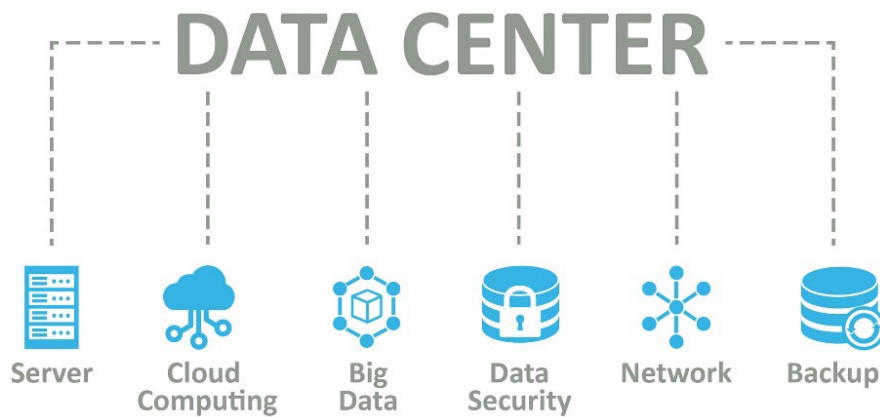


运用先进技术满足 AI 服务器日益增长的能源需求

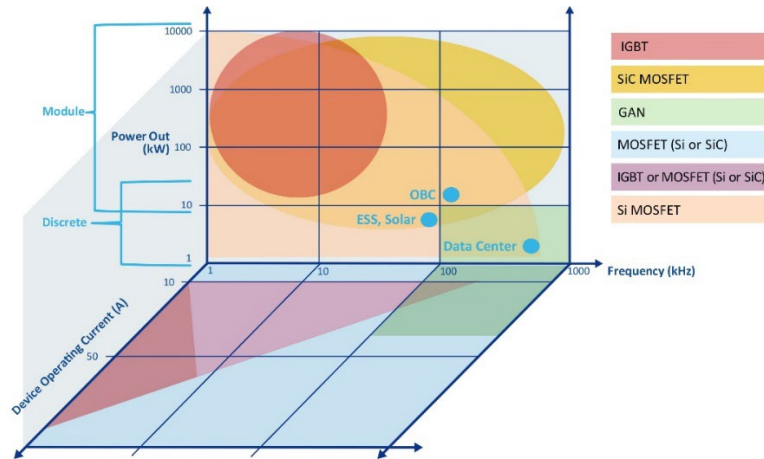
Microchip Technology Inc.
dsPIC 业务部产品营销工程师
Josue Navarro

简介

人工智能（AI）正在迅速改变各行各业，从医疗保健、金融到自动驾驶汽车和自然语言处理，推动着全方位的创新。这场革命由 AI 服务器驱动，它们提供了前所未有的计算性能。然而，AI 工作负载（包括大语言模型的广泛采用）的指数级增长导致了功耗的急剧上升，给全球数据中心带来了新的挑战。随着 AI 模型变得更加复杂以及 AI 服务器数量的增长，对强大、高效和可扩展的电力供应的需求比以往任何时候都更迫切。



现代数据中心正在不断发展以应对这些日益增长的需求。其重点正转向更高的能源效率、先进的电源管理以及集成碳化硅（SiC）等宽带隙半导体（在某些情况下包括氮化镓），以减少能量损耗。



安全性仍然是重中之重，尤其是当 **AI** 应用处理海量敏感数据时。这种环境要求采取强有力的措施，例如基于硬件的加密、安全启动机制以及警惕的实时威胁检测。为了应对 **AI** 推理和训练（特别是由大语言模型驱动的工作负载）带来的计算强度，数据中心正在采用新的供电和电压调节方法，并实施先进的热管理。可扩展性和灵活性的需求变得越来越重要，促使数据中心采用模块化基础设施并探索创新的冷却策略。随着这些趋势的持续发展，**AI** 驱动的数据中心有望在支持未来进步所需的敏捷性的同时，实现更高的效率和安全性。

为 **AI** 服务器供电的挑战

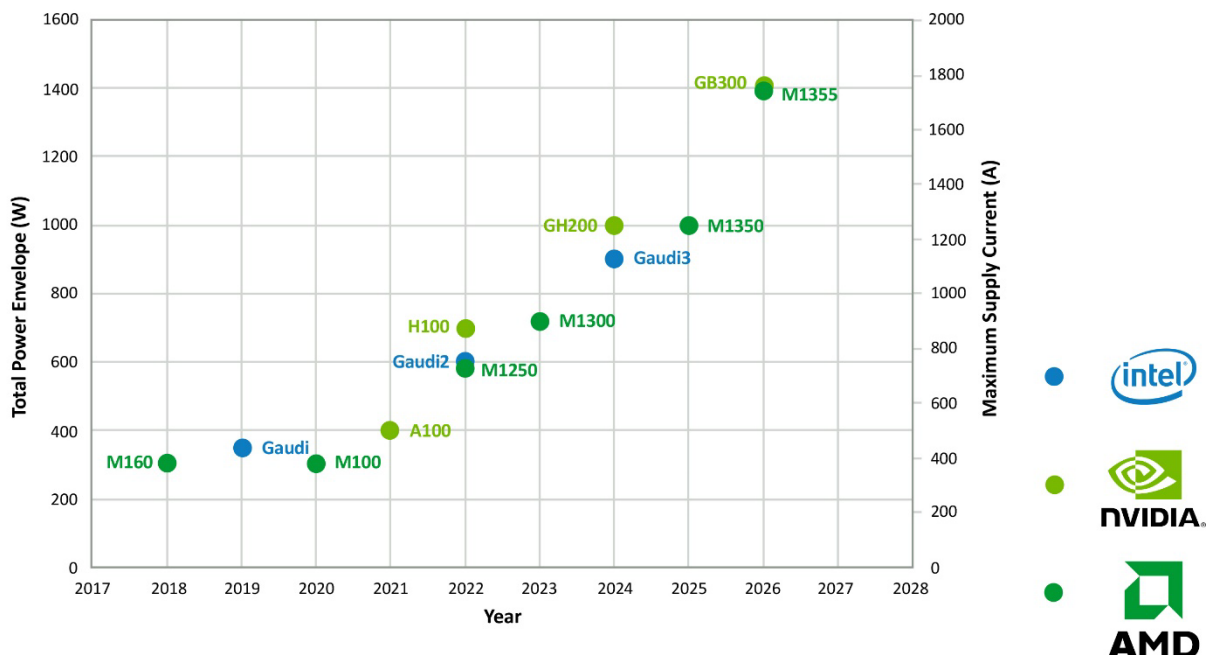


不断增长的电力和密度需求

AI 服务器是现代数据中心的支柱，驱动着深度学习、机器学习、大语言模型和实时分析等高级任务。这些服务器比传统的企业服务器需要更多的电力。随着数据中心扩展其 **AI** 能力，它们

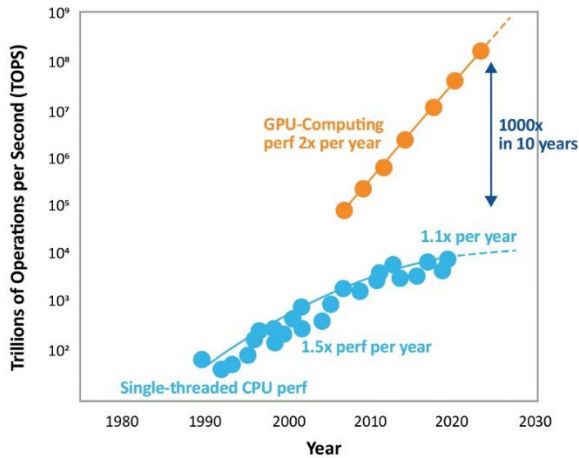
面临着在保持效率以管理成本和减少环境影响的同时，提供足够电力的挑战。数据中心的能耗激增，AI 工作负载的能耗可能是传统任务的两倍。全球数据中心的能源需求正以每年 10-15% 的速度增长，目前 AI 已占其总能耗的 10-20%。承担繁重计算负载的 AI 加速器服务器是此基础设施内的主要能源消耗者。

Increase of Load Power Demand



从传统的以 CPU 为中心的架构向 GPU 和专用加速器的转变，正推动着电力需求的持续激增。现代 AI 服务器的功耗现在是传统企业服务器的两到三倍，高性能 AI 机柜的功耗超过 50 千瓦/柜，而传统数据中心仅为 5-15 千瓦/柜。仅 GPU 每卡功耗就可达 300-700 瓦。随着 AI 工作负载的扩展，下一代数据中心必须实施先进的配电架构、高效的电压调节器和创新的冷却解决方案，以维持运行效率和可靠性。

Change to CPU/GPU Computing Performance



AI Server Power Demand Increase Power Need is Doubling

Now



Grace	300W
H200	700W
1 × Grace + 2 × Hopper GH200	1,700W

Next Generation



Grace	300W
B100	1,500W
1 × Grace + 2 × Blackwell GB200	3,300W

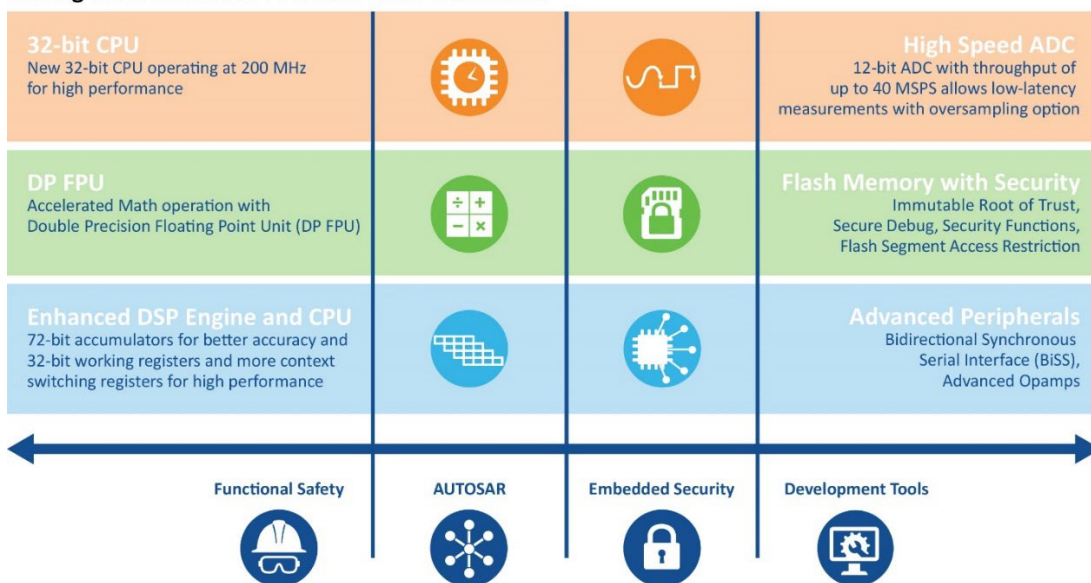
Microchip 通过其全面的高效 MOSFET、SiC FET 和智能栅极驱动器产品组合来应对这些不断增长的电力需求。这些先进的功率器件旨在提供卓越的开关性能、降低导通和开关损耗并增强热管理能力，这对于支持 AI 服务器所需的高功率密度至关重要。Microchip 的 SiC MOSFET 支持更高的开关频率，从而减小磁性元件的尺寸和重量，实现更紧凑、更高效的电源设计。智能栅极驱动器提供精确控制、集成保护功能和强大的诊断能力，确保即使在最苛刻的计算负载下也能可靠运行。这使得数据中心能够部署更强大的 AI 服务器，而不会超出功率或热极限。

效率与热管理

随着功率密度的增加，在相同或更小的物理空间内提供更多电力成为首要任务。传统的电源单元和空气冷却方法正达到其运行极限，因为更高的功率密度会产生更多的热量，并增加因效率低下而导致能量损失的风险。低效的电转换不仅会增加运营成本，还会导致更大的碳足迹，日益受到监管机构和客户的关注。有效的热管理对于防止过热、维持系统可靠性以及延长关键部件的使用寿命至关重要。数据中心必须找到最大限度地提高效率和管理散热的方法，同时最大限度地减少环境影响并实现可持续发展目标。

Microchip 的 dsPIC®数字信号控制器是其数字电源解决方案的核心，它结合了单片机的实时控制能力和数字信号处理器的高速数学处理能力。这些控制器能够实现对功率级的精确控制、快速的瞬态响应以及复杂数字控制算法的实施。例如，dsPIC33A 系列提供高时钟速度、高级脉宽调制输出和高分辨率模数转换器，从而实现对功率级的精确控制、快速的瞬态响应以及复杂数字控制算法的实施。

Driving Innovation with Precision and Performance



智能热管理和电源监控可包括温度传感器、风扇控制器和电源监控集成电路，实现对热和电气参数的实时跟踪。这些器件可以与 DSC 无缝集成，以实施闭环冷却策略、优化风扇速度并在异常情况下触发警报或关闭，确保高密度 AI 服务器的安全高效运行。

安全与数据完整性

由于 AI 服务器处理大量敏感数据，强大的安全协议至关重要。网络威胁和数据泄露风险的增加要求实施先进的基于硬件的安全措施和安全启动机制。组织还需遵守严格的行业标准，如 NIST 800-193、通用标准以及 FIPS 140-3。开放计算项目也设定了较高的安全标准，重点关注硬件信任根、固件完整性和安全启动过程。这些安全措施有助于确保 AI 数据服务器在运行前能够验证和认证硬件和软件，降低网络威胁的风险。

Microchip 将其强大的安全功能直接集成到其控制器和电源管理集成电路中。这些功能包括硬件信任根、安全启动、加密加速器以及对行业标准的支持。硬件信任根确保只有经过验证的固件和软件才能在系统上运行，而安全启动机制可防止在启动期间执行未经授权的代码。加密加速器支持基于硬件的快速加密和解密，保护静态和传输中的敏感数据。这些功能帮助数据中心满足现代安全要求，防范不断演变的网络威胁，并维护 AI 工作负载的完整性和机密性。

可扩展性与灵活性

AI 工作负载日益复杂，正推动着对 AI 数据服务器更高的可扩展性和灵活性的需求。大语言模型、实时分析和 AI 驱动应用的兴起，要求基础设施能够动态扩展以处理激增的计算需求。AI 训练集群正在迅速扩展，一些超大规模数据中心现已部署的 GPU 架构功耗超过 100 千瓦/柜。向模块化服务器架构的转变允许数据中心在无需对整个系统进行彻底改造的情况下升级和重新配置



硬件，从而降低成本并提高适应性。可组合基础设施解决方案支持根据工作负载需求动态分配计算、存储和网络等资源，确保随着新模型和应用的涌现，AI 数据服务器能够无缝扩展。

Microchip 的模块化电源管理解决方案（包括数字控制器、电源模块和参考设计）旨在支持现代 AI 工作负载所需的可扩展性和灵活性。这些解决方案可以轻松集成到模块化服务器架构中，允许数据中心根据计算需求扩展供电基础设施。Microchip 的数字控制器支持可组合基础设施，能够在工作负载演变时动态分配计算、存储和网络资源。这种方法降低了成本，提高了适应性，并确保随着新模型和应用的涌现，AI 数据服务器能够无缝扩展。

完善的开发生态系统

AI 的快速发展要求数据中心运营商和服务器制造商加快开发周期以保持竞争力。为 AI 服务器设计、验证和部署先进的电源系统非常复杂，如果没有合适的资源，可能会导致延误和成本增加。为了最大限度地缩短产品上市时间并降低设计风险，获得经过验证的参考设计、强大的开发工具和专家技术支持至关重要。

结论

AI 服务器电力需求的指数级增长给数据中心运营商和技术提供商带来了重大挑战。通过采用先进的功率器件、创新的冷却解决方案、强大的安全协议和智能数字电源管理，行业可以有效应对这些挑战。Microchip 的高效 MOSFET、精密的栅极驱动器以及具备先进 DSP 功能的 dsPIC 数字信号控制器，在实现卓越性能和能源效率的电源供应方面处于前沿地位。

通过利用先进的功率器件、数字控制器、集成的安全功能和全面的开发工具，数据中心运营商和服务器制造商能够设计出满足 AI 工作负载（包括由大语言模型驱动的工作负载）严苛要求的电源系统。这些技术实现了更高的效率和功率密度、增强的可靠性和热管理、强大的安全性、可扩展性、灵活性以及更快的开发周期。整个行业的持续创新正在帮助下一代数据中心在性能、效率 and 安全性方面达到新的基准，支持 AI 基础设施的持续演进。