

OCTC BA01—2024



开放计算标准
工作委员会
Open Compute Technology
Committee

基于标准PCIe接口的人工智能 能加速卡液冷设计

白皮书

(2024年1月)



版权保护文件

版权所有归属于该文件的发布机构，除非有其他规定，否则未经许可，此发行物及其章节不得以其他形式或任何手段进行复制、再版或使用，包括电子版，影印件，或发布在互联网及内部网络等。使用许可可于发布机构获取。

目 次

前 言	III
1 概述	1
2 发展趋势	1
3 术语	1
4 缩略语	2
5 总体描述	2
5.1 冷板式人工智能服务器组成	2
5.2 冷板式人工智能加速卡组成	3
6 技术方案	3
6.1 设计要求	3
6.2 外观要求	4
6.3 结构要求	4
6.4 热性能要求	6
6.5 可靠性要求	7
参 考 文 献	9

前 言

本文件由中移动信息技术有限公司提出。

本文件由开放计算标准工作委员会归口。

本文件起草单位：中移动信息技术有限公司，浪潮电子信息产业股份有限公司，上海壁仞智能科技有限公司，中科寒武纪科技股份有限公司，上海燧原科技有限公司，宁畅信息技术有限公司，中科可控信息产业有限公司，新华三技术有限公司，史陶比尔（杭州）精密机械电子有限公司，中航光电科技股份有限公司，京东云计算有限公司，百度在线网络技术（北京）有限公司。

本文件主要起草人：李圣义，杨洋，李金波，刘广志，罗竣峰，李明杰，周立志，邸贺亮，张政，黄乾明，王辉，朱重兴，孙鹏，李壮，谢泽邦，王周杰，廖嵬，聂超，高亦廷，梅敬青，王思善，王海岩，韩学磊，李伟祖，王伟，李鹏，谢地，董永申，李进宝，王贵林，蔡岳霖，金跃红，郑建武，高从文，刘玲，董少杰，郗卓宁，傅浩杰。

基于标准 PCIe 接口的人工智能加速卡液冷设计指南

1 概述

冷板式人工智能加速卡在结构设计、流体快插接头、热性能设计等方面无统一设计要求，服务器针对不同厂商的冷板式人工智能加速卡需要进行结构和散热适配，耗费大量人力、物力。本技术白皮书主要制定标准PCIe接口形态的冷板式人工智能加速卡在结构设计、冷板设计、流体快插接头选型、热性能设计、可靠性设计方面的要求，用于指导冷板式人工智能加速卡的设计。通过本白皮书提供的设计指导，降低服务器和冷板式人工智能加速卡之间的适配难度，降低适配的成本和人力投入。同时降低部署和运维难度，方便终端客户根据业务场景灵活选择相应产品，加速液冷式人工智能加速卡和服务器产品上市。此外，通过统一的设计指导，在冷板本体、流体快插接头等关键组件上产生规模化效益，获取低成本液冷解决方案，推动冷板式液冷技术的普及。

2 发展趋势

以AIGC为代表的人工智能大模型的发展，带来了算力需求的指数式提升。AI芯片需要通过提升工作频率或增加运算单元来获得更强的算力，带来人工智能加速卡功耗的不断提升。当前人工智能加速卡单卡功耗已达到350W至600W量级，带来服务器散热功耗的显著提升，在给服务器散热设计带来挑战的同时，也给建设绿色数据中心带来了很大的挑战。液冷技术成为解决人工智能服务器高功耗散热问题、降低数据中心PUE的关键，其中冷板式液冷技术是当前解决人工智能加速卡高功耗问题的主流方案。目前标准PCIe接口形态的人工智能加速卡仍然是业界主流，国内外部分厂商已经推出了相应的冷板式人工智能加速卡，但均是按照自身产品形态特征进行设计，未考虑同行业其他厂商情况，不能作为标准化设计。

同时，虽然冷板式液冷散热技术在CPU、内存、扣板式人工智能加速卡上已经取得了成功应用，但在标准PCIe接口形态的人工智能加速卡上的应用仍然很少，业界迫切需要一个针对标准PCIe接口形态人工智能加速卡的液冷设计指导，来推动冷板式散热技术在人工智能加速卡应用的普及。

3 术语

3.1

冷板式液冷 cold plate liquid cooling

通过冷板（通常为铜铝等导热金属构成的封闭腔体）将发热器件的热量间接传递给封闭在循环管路中的冷却液体，通过冷却液体将热量带走的一种实现形式。

3.2

冷板式人工智能加速卡 cold plate liquid cooling artificial intelligence accelerator card

专为人工智能应用设计，通过冷板进行散热的硬件加速器，可用于提升机器学习和深度学习的速度和性能，从而更好地满足人工智能应用的需求。

3.3

冷板式人工智能服务器 cold plate liquid cooling artificial intelligence server

配备了冷板式人工智能加速卡的服务器，可为人工智能应用提供高效能计算处理能力。

3.4

机柜冷却工质供回歧管 rack coolant manifold

用于向机柜内各液冷冷板分配冷却工质的装置。

3.5

流体快插接头 hydraulic quick disconnect

一种包含插头和插座、且插头和插座都带流体截断功能的快速插拔组件。

4 缩略语

下列缩略语适用于本文件。

AI 人工智能（Artificial Intelligence）

AIGC 生成式人工智能（Artificial Intelligence Generated Content）

ASIC 专用集成电路芯片（Application Specific Integrated Circuit）

CEM 板卡机电（Card Electromechanical）

CPU 中央处理器（Central Processing Unit）

EMC 电磁兼容性（Electromagnetic Compatibility）

FPGA 现场可编程门阵列（Field Programmable Gate Array）

GPGPU 通用图像处理器（General-purpose computing on graphics processing units）

PCIe 高速外围组件互连总线（Peripheral Component Interconnect Express）

PUE 数据中心电能利用率（Power Usage Effectiveness）

RCM 机柜冷却工质供回歧管（Rack Coolant Manifold）

VR 电压调节器（Voltage Regulator）

5 总体描述

5.1 冷板式人工智能服务器组成

人工智能服务器组成如图1所示，主要包括通用计算子系统、异构加速子系统、存储子系统、互联子系统、监控子系统、供电子系统、结构及散热子系统、输入输出设备等组成。其中异构加速子系统配备了面向人工智能应用的加速器，例如GPGPU卡、AI ASIC卡、FPGA卡等，可为人工智能应用提供高效能计算处理能力。

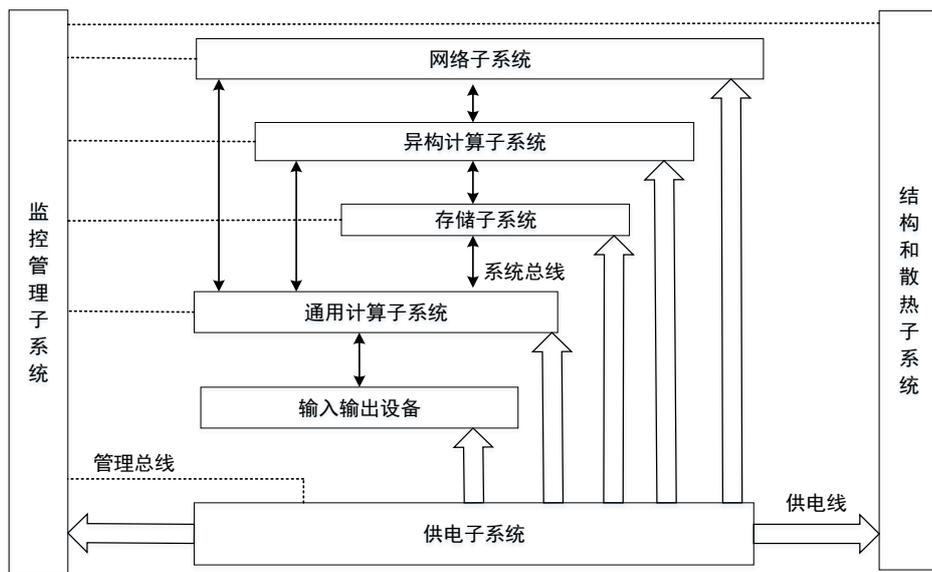


图1 人工智能服务器组成

人工智能服务器基于散热子系统实现方式的不同,可分为风冷式人工智能服务器和冷板式人工智能服务器,其中冷板式人工智能服务器是本文研究的重点。

冷板式人工智能服务器CPU和人工智能加速卡应采用冷板进行散热,内存等其它高功耗部件宜采用冷板散热,以进一步提升液冷散热的占比,降低数据中心PUE。设计上需要应减少机箱内的管路,并采用漏液检测机制,以提升服务器的可靠性。

5.2 冷板式人工智能加速卡组成

冷板式液冷加速卡由人工智能加速板卡、液冷板、加速卡外壳结构件组成。

冷板覆盖加速卡上的主芯片及其他发热元件(包括但不限于VR、显存等),提高板卡封装狭小空间内人工智能芯片散热效率。加速卡对外提供一对流体快插接头(两个公头),用于连接服务器内部集/分水结构或RCM。

冷板式人工智能加速卡能够有效避免局部热点,降低服务器系统风扇功耗,降低数据中心整体PUE,是未来绿色数据中心发展的必然趋势。

6 技术方案

本技术方案旨在针对标准PCIe接口的人工智能加速卡冷板设计的外观、结构、液冷快插接头、热性能、可靠性方面进行定义,以解决冷板式人工智能加速卡和服务器适配问题。

6.1 设计要求

人工智能加速卡液冷设计要求如下:

- 应根据AI芯片的型号尺寸和发热特点及电子信息设备的内部结构进行设计,以获得更好的换热效率,在满足芯片整个使用周期内的壳温要求下,应优化流道设计,减小冷板模块的流阻;
- 应保障满足AI芯片插座的载荷及其他结构性要求;
- 应考虑配管位置及方向,液体进出口位置,避免与电子信息设备产生干涉;

- d) 冷板基板和流道宜采用铜或铝合金材质，一个系统中冷却工质直接接触的部件不应有两种电极电位差较大的金属；
- e) 冷却工质的选用应考虑与二次侧循环回路中所有直接接触的固体表面材质间的相容性；
- f) 漏液检测装置的泄漏量感应基准应不高于0.5ml；
- g) 应符合芯片对散热器重量的要求；
- h) 应考虑冷板的安装及拆卸顺序，满足芯片的操作规范；
- i) 应满足芯片的扣合力技术要求，安装拆除后散热基板底面满足平面度技术要求。

6.2 外观要求

人工智能加速卡外观要求如下：

- a) 冷板散热盖板应光滑，不应变形，散热基板底部表面不应有裂纹，划痕，变形、污点等缺陷；
- b) 冷板散热基板散热面形状宜为方形或八边形；
- c) 固定模块表面锐边倒钝，无毛刺，外表面无划痕、脏污，明显色差和花斑、裂缝、变形等缺陷，涂覆层无起泡、堆积、龟裂和脱落现象；
- d) 接口应无毛刺、划痕、变形等缺陷，并与配管顺畅接合；
- e) 配管管内应无脏污，无毛边披锋，无破损。

6.3 结构要求

6.3.1 加速卡结构尺寸要求

标准PCIe接口的人工智能加速卡挡片宜不开孔，冷板全覆盖加速卡进行全液冷设计。加速卡对外液冷接口需要考虑和服务器系统或RCM的对接方式，有利于加速卡和外界液冷系统对接。板卡尺寸符合PCIe CEM规范，为单槽或者双槽、全高形态，卡长（不含快插接头）宜不大于266.7mm。加速卡的进出水口快插接头可根据应用场景放置于加速卡挡片侧或者加速卡尾端。

6.3.1.1 当流体快插接头位于挡片侧，结构尺寸设计要求如图2所示。

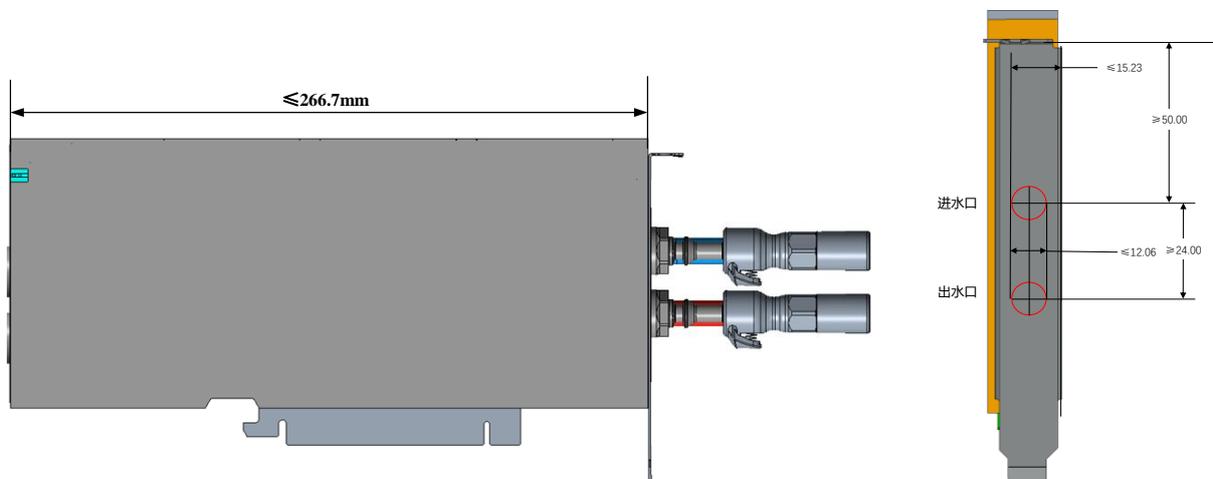


图2 结构设计要求（流体快插接头位于挡片侧）

人工智能加速卡需要插入服务器才能工作，因服务器后窗针对液冷卡有限制，需要进出水口能满足下列要求：

- 如图2所示，加速卡流体快插接头位于挡片侧，进水口位于出水口上方。快插接头和冷板连接的地方，需要做凸台，凸台高度与机箱后窗结构相关，宜高度不小于3mm，以方便密封和在服务器侧的插拔；
- 加速卡公头位于加速卡上，两个快插接头公头中心点之间距离大于等于24mm，以满足快插接头母头能支持正常的插拔操作，且无结构干涉；
- 加速卡快插接头母头最大直径19.5mm，宜小于18mm，以满足两张相邻的加速卡之间快插接头正常的插拔操作；
- 进水口快插接头中心点距离挡片最上方距离大于等于50mm，以避免加速卡插入后挡片和板卡结构干涉，确保快加速卡插接头可以从服务器后窗推出；
- 快插接头的外径需要小于机箱后窗开口的宽度，最大不超过12.06mm；
- 快插接头的左边缘距离挡片右边缘应小于等于15.23mm。

6.3.1.2 当进出水快插接头位于尾部时，设计要求如图3所示。

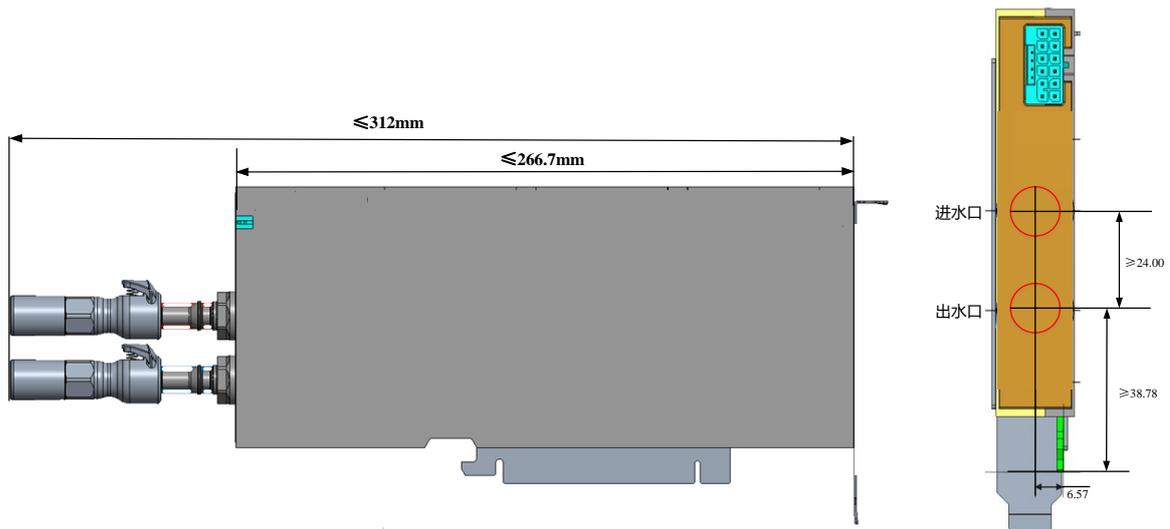


图3 结构设计要求（流体快插接头位于加速卡尾部）

- 如图3所示，加速卡流体快插接头位于尾部，进水口位于出水口上方；为了节省空间，快插接头和冷板连接的地方不宜有凸台；
- 加速卡公头位于加速卡上，两个快插接头公头中心点之间距离大于等于24mm，以满足快插接头母头能支持正常的插拔操作，且无结构干涉；
- 加速卡快插接头母头最大直径19.5mm，宜小于18mm，以满足两张相邻的加速卡之间快插接头正常的插拔操作；
- 出水口快插接头中心点距离加速卡金手指下边缘大于等于38.78mm；
- 为了便于服务器机箱设计，流体快拆头位于尾部时，宜加速卡本体+快插接头静态插合尺寸不超过312mm。

6.3.2 加速卡冷板设计要求

加速卡冷板的设计要求如下：

- 冷板的材料需要考虑导热性及与冷却液的化学兼容性，例如可选择紫铜。冷板对板卡上热源的全覆盖；
- 固定方式宜使用4个弹簧螺丝固定的方式进行固定；
- 主芯片和冷板之间的压力应保证满足散热性能需求；
- 冷板散热基板底部和主芯片接触区域要求光滑，且平面度不大于0.05mm，粗糙度Ra不大于1.6um。

6.3.3 流体快插接头要求

人工智能加速卡流体快插接头选型主要考虑结构兼容性，需要保证在加速板卡侧面正常连接冷板和封装，同时组装在服务器上不发生结构干涉。为了满足快插接头位于人工智能加速卡挡片侧或者尾部两种不同场景下的结构要求，流体快插接头尺寸要求见表1。

表1 流体快插接头规格要求

流体快插接头参数	规格要求
公头最大直径	≤12.06mm
母头最大直径	≤19.5mm，推荐不大于18mm
静态插合尺寸	≤45mm

静态插合尺寸是指快插接头公头和母头插合后，公头橡胶圈和母头橡胶圈之间的尺寸，如图4所示。



图4 流体快插接头静态插合尺寸示意图

流体快插接头参考型号见表2。

表2 流体快插接头参考型号

厂商	参考型号
中航光电科技股份有限公司	TSC/D-4
史陶比尔（杭州）精密机械电子有限公司	紧凑版 SCG03/IC

6.4 热性能要求

6.4.1 人工智能加速卡热性能参数

冷板式人工智能加速卡热性能参数及建议值见表3。

表3 热性能参数要求

热性能参数	建议值
加速卡热设计功耗	单宽卡: $\leq 450\text{W}$ 双宽卡: $\leq 600\text{W}$
加速卡进水温度	$30^{\circ}\text{C}\sim 55^{\circ}\text{C}$ (最低温度为参考值, 宜高于机房凝露温度 3°C)
加速卡最大出水温度	65°C (参考值)
进出水口温升	进出水口温升宜 $6^{\circ}\text{C}\sim 10^{\circ}\text{C}$ 左右;
进出水口最大压降 @额定流量	进水口到出水口总压降 (含快插接头) $\leq 45\text{kPa}$ (按最大 600W 功耗计算)
翅片流速	经过翅片的最大流速 $\leq 2\text{m/s}$
冷却液颗粒物要求	$\leq 50\mu\text{m}$, 参考标准 YD/T 3982-2021《数据中心液冷系统冷却液体技术要求和测试方法》: 液体中固态颗粒物含量, $< 0.5\mu\text{m}$ 颗粒不超过 25 颗/ml, $0.5\sim 1\mu\text{m}$ 颗粒不超过 5 颗/ml, $1\mu\text{m}\sim 3\mu\text{m}$ 颗粒不超过 3 颗/ml, 没有直径超过 $50\mu\text{m}$ 金属颗粒。

6.4.2 人工智能加速卡润湿材料

润湿材料选型要求见表4。

表4 润湿材料选型要求

元件	建议材料
冷却工质	需要考虑冷却工质热物性、环保特性、安全特性、环境适应性、工质与管路材料兼容性、工质成本和可获取性等。IT 设备推荐使用含缓蚀剂、杀菌剂等药剂的水、乙二醇/丙二醇水溶液等。冷却液标准参考 YD/T 3982-2021《数据中心液冷系统冷却液体技术要求和测试方法》。
冷板	冷板材质优选: 紫铜、铝合金 (需做好耐腐蚀和绝缘要求)。
快插接头	和冷却液具有化学兼容性。 不锈钢、铝合金 (仅限搭配铝合金冷板)、黄铜 (表面需要做好镀层, 把黄铜和液体隔离, 冷却液不推荐使用去离子水)。

6.5 可靠性要求

6.5.1 漏液检测要求

服务器系统设计宜有漏液检测装置 (例如采用漏液感应线), 每个业务单板提供一个监控连接器接口 (服务器系统来定义), 检测范围覆盖冷板易泄漏点, 比如: 焊接部位, 管链接位置等。

6.5.2 环境适用性

工作、贮存中温度、湿度、EMC、压力要求见表5。

表5 环境适用性要求要求

参数	建议值
冷板最大工作压力	350kPa, 无液体泄露, 无变形
冷板最小爆破压力	最大工作压力的 3 倍, 无液体泄露, 可变形
温度	工作: 0°C~40°C 贮存: -45°C~85°C
湿度	工作: 8%RH~90%RH (无冷凝) 贮存: 10%RH~93%RH (无冷凝)
大气压	86kPa~106kPa
EMC CLASS A	搭配服务器去获取相关认证, 实际设计参考国家标准 GB/T 17625.9-2016 《电磁兼容 限值 低压电气设施上的信号传输 发射电平、频段和电磁骚扰电平》。

参 考 文 献

- [1] GB/T 17625.9-2016 《电磁兼容 限值 低压电气设施上的信号传输 发射电平、频段和电磁骚扰电平》
 - [2] YD/T 3982-2021 《数据中心液冷系统冷却液体技术要求和测试方法》
 - [3] T/CESA 1249.1-2023 《服务器及存储设备用液冷装置技术规范》
-