

团体标准《液冷式人工智能加速卡设计技术要求》(征求意见稿)编制说明

一、工作简况

1.1、仟务来源

本标准的编制任务来源于中国电子工业标准化技术协会2023年第九批团体标准项目计划,正式批复团体标准《液冷式人工智能加速卡设计技术要求》(项目号: CESA-2023-086)立项。技术归单位是中国电子工业标准化技术协会开放计算标准工作委员会。

1.2、起草单位

本标准由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牵头,参与单位包括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、中航光电科技股份有限公司、史陶比尔(杭州)精密机械电子有限公司、京东云计算有限公司、百度在线网络技术(北京)有限公司等。根据标准编写组的任务分工,牵头单位负责标准相关资料的搜集和调研、标准框架编制、标准内容起草、反馈意见整理等工作。

1.3、编制过程

主要编制过程总结如下:

(一) 标准预研

为推动液冷式人工智能加速卡部件行业应用,提高液冷式人工智能加速卡产品质量,降低应用门槛及制造成本,进而有效地推广液冷式人工智能服务器在数据中心的应用,从而应对日益增加的散热瓶颈和人工智能算力需求,有效降低数据中心能耗,2023年4月中移动信息技术有限公司联合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对液冷式人工智能加速卡共性需求进行了分析,初步确定了液冷式人工智能加速卡设计技术要求的标准草案结构和主要内容。2023年4月至2023年9



月,标准工作组定期组织9次标准讨论会,对标准制定的必要性,可行性,目的 意义,拟解决的问题,标准范围进行了充分讨论,最终取得技术共识。

2022年9月,中移动信息技术有限公司联合浪潮电子信息产业股份有限公司, 共同向冷板制造商,流体快插接头制造商,人工智能芯片厂商,人工智能服务器 集成商,人工智能服务器设备使用方发起参编邀请,各意向参编单位对于该标准 制定的必要性,可行性,目的意义,拟解决的问题表示一致认同并共同推进标准 立项。

(二) 标准立项

2023年10月,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牵头,联合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、中航光电科技股份有限公司、史陶比尔(杭州)精密机械电子有限公司、京东云计算有限公司、百度在线网络技术(北京)有限公司作为共同发起方申请立项,并通过评审,成为协会正式标准制定项目。

(三) 标准编制

2023年10月,标准工作组完成了面向工作组内成员单位的标准意见征集,并组织召开了标准启动会,对标准制定的背景、参编单位构成,标准推动计划进行了介绍。同时对标准范围等相关意见进行了讨论。

2023年11月-2024年2月,标准工作组组织召开五次讨论会,重点对人工智能加速卡的设计、外观、结构、热性能、环境适应性、多卡互连的设计要求等相关意见进行了讨论。

2024年3月21日,标准工作组组织召开讨论会,重点对设计、外观、结构、 承压、热性能、环境适应性试验方法的相关意见进行了讨论。

2024年4月11日,标准工作组组织召开讨论会,重点对冲击适用性测试、运输包装件跌落测试、气密性检测、高低温试验、高温高湿试验、测试的适用性等相关意见进行了讨论。结合历次标准工作组会议上对标准意见的处理情况,牵头单位完成对标准草案的刷新,形成征求意见稿。



2024年4月25日,标准工作组组织召开讨论会,对气密性试验方法和判据进行了讨论;标准工作组完成对征求意见稿的确认,并对当前征求意见稿中是否涉及知识产权信息进行了检索和反馈。

2024年5月16日,标准工作组组织召开讨论会,重点结合实际试验报告,对 气密性试验方法和判据再次进行了讨论,并刷新征求意见稿。

2024年5月30日,标准工作组完成对征求意见稿刷新的确认。

二、标准编制原则和确定主要内容的论据及解决的主要问题

2.1、编制原则

在标准编制过程中考虑到液冷式人工智能加速卡的特点, 遵循了以下五方面的原则。

- a) 符合性: 遵循国家法律、法规等相关规定,制定过程严格按照程序 执行。
- b) 先进性:本标准制定过程中充分考虑了人工智能芯片和冷板的技术 发展,并在方面保持了一定的前瞻性。
- c) 适用性: 当前国内的人工智能加速卡仍以风冷散热为主,没有液冷式人工智能加速卡的设计标准,一定程度上影响了液冷式散热技术在人工智能加速卡领域的普及。本标准规范液冷式人工智能加速卡的构成及技术规格,指导液冷式人工智能加速卡的设计、制造、采购和评价活动,解决由于标准不统一所导致的组件高度定制,产业生态萧索,生态链不完整等问题。为液冷式人工智能加速卡产品的品质保证提供统一的衡量标准,为制造商生产液冷式人工智能服务器、为最终用户的产品选型、为液冷式人工智能服务器产品市场准入提供依据。
- d) 中立性:本文件制定过程中编制组成员单位对标准文本进行了充分 讨论。
- e) 科学性:本标准主要解决市场中由于标准不统一所带来的生态构建和产业化推广问题,给液冷式人工智能加速卡的设计提供了标准支撑,同时有助于液冷式人工智能服务器生态构建和完善,实现大规模部署,进而带来绿色节能、快速交付、便捷运维等收益。本标准



与国内行业标准协调一致,参考国内团体标准,制定液冷式人工智能加速卡设计技术要求。本文件技术指标科学合理并可验证,试验方法具有稳定性。

2.2、主要内容

本标准规定了采用冷板式液冷散热的标准PCIe接口形态的人工智能加速卡设计要求和技术要求,并描述了相关技术要求的试验方法。本标准适用于采用冷板式液冷散热的标准PCIe接口形态的人工智能加速卡/服务器的设计、制造和测评,其它结构形态的插卡式人工智能加速卡的冷板式液冷散热设计也可参考。

设计要求部分,主要规定了冷板设计、漏液检测装置设计、多卡互联设计等加速卡关键技术设计要求。相关要求满足整机企业及下游最终客户使用要求,提高加速卡的适配及应用效率。技术要求部分,主要规定了外观要求、结构要求、热性能要、环境适用性要求及电磁兼容性要求,满足服务器整机对加速卡部门的基本要求,保证加速卡装机后的服务器整机的基本性能。技术内容经冷板制造商,流体快插接头制造商,人工智能芯片厂商,服务器厂商讨论及验证。试验方法主要规定了技术要求项目的相关试验方案,统一加速卡上下游企业相关测试验证及验收基准。

2.3、解决的主要问题

当前基于标准PCIe接口的人工智能加速卡单卡功耗目前已达600W及以上,单卡功耗的提升,带来服务器散热子系统功耗的显著提升。节能减排、降低PUE、建立绿色数据中心也是发展趋势所在。液冷成为解决AI服务器系统高功耗散热问题和降低数据中心PUE的关键措施,其中冷板式液冷技术日渐成熟,成为解决人工智能加速卡高功耗问题的主流方案。当前液冷式人工智能加速卡在硬件接口、结构形态、快拆头等方面无统一要求,适配不同厂商液冷卡,服务器内部需要进行不同的结构设计,加速卡和整机适配需要耗费大量人力、物力。不同的冷板设计会对加速卡和系统的适配开发造成非常大的影响。

基于以上原因,有必要定义统一的液冷式人工智能加速卡接口形态,降低服务器适配难度,规范液冷数据中心设计。本标准主要制定标准PCIe接口的液冷式人工智能加速卡在结构、冷板、接头方面的设计要求。通过制定统一的设计指导,统一加速卡PCBA结构、冷板本体、快拆头等液冷式散热人工智能加速卡关键组件,



通过规模化效益获得低成本液冷式人工智能加速卡解决方案;降低液冷式散热人工智能加速卡适配难度和周期,加速人工智能芯片上市和迭代周期;方便终端客户灵活选择适配业务场景的解决方案,降低部署和运维难度,提高部署效率;降低服务器针对不同厂商液冷式人工智能加速卡进行适配的成本投入和人力投入。

三、主要试验[或验证]情况分析

在标准起草过程中,标准工作组对标准规定的液冷式人工智能加速卡设计技术要求和试验方法征集了冷板制造商,流体快插接头制造商,人工智能芯片厂商,服务器厂商的意见,相关技术方案经讨论确定相关技术要求。标准方案经过浪潮信息、宁畅整机厂家适配验证,并比对及满足了移动、百度、京东等下游用户单位的应用需求。

本标准充分考虑科技发展及应用,符合市场及用户发展需求;本标准规定的 技术要求和试验方法是合理的,切实给出符合人工智能加速卡产业实际需要的要求。

四、知识产权情况说明

本标准不涉及国内外专利等知识产权问题。

五、产业化情况、推广应用论证和预期达到的经济效果

标准发布后,牵头及参编单位将在国内重量级会议进行推广,人工智能加速卡和芯片厂商基于本文件可作为冷板、流体快插接头选型与使用的依据,部件制造厂商基于本文件进行冷板、流体快插接头设计与测试,服务器厂商、终端应用厂商可将本文件作为液冷人工智能加速卡选型与液冷数据中心建设的参考规范。

六、转化国际标准和国外先进标准情况

目前国际标准空白,本标准为自主制定,未采用国际标准,本标准制定水平符合当前液冷式人工智能加速卡设计的技术要求。

七、与现行相关法律、法规、规章及相关标准的协调性

该标准与我国的现行法律、法规和标准是协调一致。目前标准 PCIe 接口的液冷式人工智能加速卡设计无相关国家或行业标准。目前已有 1 项与服务器冷板设计的标准 T/CESA 1249.1-2023《服务器及存储设备用液冷装置技术规范》以及 1 项与数据中心电子信息设备液冷相关的行业标准 YD/T 3980-2021《数据中心冷板式液冷服务器系统技术要求和测试方法》。相关标准主要概述了数据中心冷板



式液冷系统设计、液冷冷却液相关方面的技术要求与测试方法,适用于液冷系统的设计、施工、部署、运维、测试等环节的技术指导。但对于人工智能加速卡冷板的结构、热性能、快拆头等规格标定、设计、生产、测试和评价标准,内容较少。

八、重大分歧意见的处理经过和依据

无

九、贯彻标准的要求和措施建议

建议列为推荐性标准,在标准发布后尽快组织标准宣贯、试验验证。

十、替代或废止现行相关标准的建议

无需要替代或废止的现行相关标准。

十一、其它应予说明的事项

无

《液冷式人工智能加速卡设计技术要求》 团体标准编制起草组 2024-05-30