

证券简称：寒武纪

证券代码：688256

## 中科寒武纪科技股份有限公司

### 投资者关系活动记录表

编号：2021-009

<b>投资者关系活动类别</b>	<input type="checkbox"/> 特定对象调研 <input checked="" type="checkbox"/> 分析师会议 <input type="checkbox"/> 媒体采访 <input type="checkbox"/> 业绩说明会 <input type="checkbox"/> 新闻发布会 <input type="checkbox"/> 路演活动 <input type="checkbox"/> 现场参观 <input checked="" type="checkbox"/> 电话会议 <input checked="" type="checkbox"/> 其他 <u>券商策略会</u>
<b>参与单位名称</b>	<p>（排名不分先后，按字母顺序排列）Aspex Management、博时基金、创金合信基金、大成基金、东方阿尔法、东方证券、方圆基金、富国基金、Fuh Hwa、富唐资管、Fountain Capital、GF Asset Management HK、工银瑞信基金、固禾资产、广东宝新资管、广发证券、观富资管、国金证券、航长投资、Hao Fund、Hel Ved Capital、恒健远志、恒生前海基金、河清资本、HSZ Group、华安基金、华商基金、华西证券、汇丰晋信基金、混沌道然资产、Invesco、加拿大鲍尔、嘉实基金、景泰利丰、聚联汇投资、开思基金、澜和基金、马可波罗基金、民生证券、Pinpoint Asset Management、Point 72 Asset Management、Prudence Investment Management、浦发银行、前海互兴、前海开源、前海联合基金、润晖投资、山石基金、尚诚投资、申港证券、神农投资、申万宏源、盛曦投资、Sherwood Capital、首创证券、泰达宏利基金、太平保险资管、太平资产、唐融资产、Trikon AM、拓璞基金、望正资本、微明恒远、文渊资本、谢诺辰阳、信诚基金、新同方投资、鑫元基金、幸福时光基金、兴业证券、易方资本、一鸣投资、英大国际信托、宜信投资、永赢基金、远策投资、Yunqi Path、智诚海威、中金公司、中金</p>

	证券、中泰国际、中信保诚基金、中信证金、中信证券、中信证券资管、珠海万方资产。
时间	2021年11月4日、2021年11月11日、2021年11月17日、2021年11月18日、2021年11月30日
形式	<input type="checkbox"/> 现场 <input checked="" type="checkbox"/> 网络会议 <input checked="" type="checkbox"/> 电话会议
地点	北京、上海
上市公司接待人员姓名	<p>董事长、总经理：陈天石</p> <p>副总经理、首席技术官：梁军</p> <p>董事、副总经理、财务负责人、董事会秘书：叶淦尹</p> <p>证券事务代表：童剑锋</p> <p>（注：除叶淦尹女士参加全部场次外，其他人员参与部分场次）</p>
投资者关系活动主要内容介绍	<p><b>一、请问公司的核心技术优势是如何体现在或者转化为思元370系列产品的核心优势或者亮点？</b></p> <p>答：（1）芯片算力和实测性能、能效提升：凭借7nm制程工艺和寒武纪最新智能芯片架构MLUarch03，思元370峰值算力可达256TOPS(INT8)，是寒武纪第二代产品思元270算力的2倍。同时，思元370芯片支持LPDDR5内存，高带宽且低功耗，可在板卡有限的功耗范围内给AI芯片分配更多的能源，输出更高的算力。相较于峰值算力的提升，思元370在实测性能和能效方面也具有一定优势。以ResNet-50为例，MLU370-S4加速卡（半高半长）实测性能为同尺寸主流GPU的2倍；MLU370-X4加速卡（全高全长）实测性能与同尺寸主流GPU相当，能效则大幅领先。</p> <p>（2）产品配置灵活，满足不同应用场景需求：寒武纪首次采用chiplet技术，通过不同芯粒组合规格多样化的产品，为用户提供适用不同场景的高性价比AI芯片。</p> <p>（3）推动“云边端一体、训推一体、软硬件协同”的统一智能生态建设：为了加快用户端到端业务落地的速度，减少</p>

模型训练研发到模型部署之间的繁琐流程,寒武纪的统一基础软件平台 Cambricon Neuware 整合了训练和推理的全部底层软件栈,将 MagicMind 和深度学习框架 Tensorflow、Pytorch 深度融合,实现训推一体,有效提升开发部署效率,降低用户成本。

二、公司新发布的思元 370 首次采用了 **chiplet** 技术,请简要介绍下该技术。这项技术对于提升产品性能或者业务推广有何助益?

答:思元 370 是寒武纪首款采用 chiplet (芯粒) 技术的云端智能芯片,在一颗芯片中封装 2 颗 AI 计算芯粒(MLU-Die),每一个 MLU-Die 具备独立的 AI 计算单元、内存、IO 以及 MLU-Fabric 控制和接口,通过 MLU-Fabric 保证两个 MLU-Die 间的高速通讯。得益于芯粒技术,思元 370 可以通过不同 MLU-Die 组合规格多样化的产品,为客户提供适用不同场景的高性价比 AI 芯片。

三、公司新品发布中提到思元 370 是基于寒武纪最新智能芯片架构 MLUarch03,此前公司半年报中也披露了公司第四代智能处理器微架构 MLUarch03 的研发进展。是否可以理解为公司自研的处理器微架构对于智能芯片的设计和研发有着至关重要的意义和作用?

答:智能芯片研发需要全面掌握核心芯片与系统软件的大量关键技术,技术难度大、涉及方向广,是一个极端复杂的系统工程,其中处理器微架构与指令集两大类技术属于最底层的核心技术。公司自成立以来长期开展智能芯片架构研发工作,迄今已自主研发了四代智能处理器微架构 (MLUarch00、MLUarch01、MLUarch02 和 MLUarch03),是国内外在该技术方向积累最深厚的企业之一。

由公司自主研发的第四代智能处理器微架构

(MLUarch03), 拥有新一代张量运算单元, 内置 Supercharger 模块大幅提升各类卷积效率; 采用全新的多算子硬件融合技术, 在软件融合的基础上大幅减少算子执行时间。MLUarch03 芯片架构在计算单元、多核处理、控制指令、计算指令、访存能效和可编译性等多方面均实现了优化和提升。在寒武纪新一代人工智能芯片架构 MLUarch03 的加持下, 思元 370 峰值算力提升, 最高可达 256TOPS(INT8), 而且思元 370 实测性能和能效方面的表现也较为优秀。

**四、公司推出思元 370 的同时, 还发布了两款基于思元 370 的加速卡 MLU370-S4 和 MLU370-X4。请介绍下这两款 AI 加速卡的产品性能和市场竞争力如何?**

答: MLU370-S4 智能加速卡搭载思元 370 芯片, 采用 7nm 制程工艺和寒武纪新一代人工智能芯片架构 MLUarch03, 板卡功耗为 75w。该款加速卡体积小、能效出色, 可在服务器中实现高密度部署。具体而言, (1) 实测性能方面, MLU370-S4 加速卡的性能平均接近市场主流 70W GPU 的 2 倍; (2) 能效方面, 相较于同尺寸市场主流 GPU, MLU370-S4 处理相同 AI 任务的用电量可减少 50% 以上, 将有力地帮助用户实现“双碳”目标。此外, MLU370-S4 加速卡在解码方面具有较强竞争力, 相较于同尺寸 GPU, 可提供 3 倍的解码能力和 1.5 倍的编码能力。

MLU370-X4 智能加速卡同样搭载思元 370 芯片, 采用 7nm 制程工艺和寒武纪新一代人工智能芯片架构 MLUarch03, 为单槽位 150w 全尺寸加速卡。该加速卡的优势表现为高性能, 可提供 256TOPS(INT8)推理算力和 24TFLOPS(FP32)训练算力, 同时提供 FP16、BF16 等多种训练精度, 配合全新 Cambricon Neuware 软件栈, 可充分满足推训一体 AI 任务需求。

五、公司本次对寒武纪的统一基础软件平台也进行了全面升级，整合了训练和推理的全部底层软件栈，是否意味着公司的软件栈已在通用性方面取得长足的进步和完善？

答：公司为云边端全系列智能芯片与处理器产品提供统一的平台级基础系统软件 Cambricon Neuware，并通过持续研发和升级，以适配新的芯片。本次全新升级的统一基础软件平台 Cambricon Neuware 整合了训练和推理的全部底层软件栈，新增推理加速引擎 MagicMind，将 MagicMind 和深度学习框架 Tensorflow、Pytorch 深度融合，实现训推一体。在寒武纪全系列计算平台上，从云端到边缘端，用户均可以无缝地完成从模型训练到推理部署的全部流程，进行灵活的训练推理业务混布和潮汐式的业务切换，可快速响应业务变化，提升开发部署的效率，降低用户的学习成本、开发成本和运营成本。

在通用性方面，Cambricon Neuware 支持 FP32、FP16 混合精度、BF16 和自适应精度训练等多种训练方式并提供灵活高效的训练工具，高性能算子库已完整覆盖视觉、语音、自然语言处理和搜索推荐等典型深度学习应用，可满足用户对于算子覆盖率以及模型精度的需求。

六、公司全新打造的 MagicMind 是否提升了 Cambricon Neuware 软件栈的性能和竞争力？

答：MagicMind 是公司全新打造的推理加速引擎。在 MLU、GPU、CPU 训练好的算法模型上，借助 MagicMind，客户仅需投入极少的开发成本，即可将推理业务部署到寒武纪全系列产品上，并获得具有竞争力的性能。

MagicMind 的优势不仅在于可以提供较高的性能、可靠的精度以及简洁的编程接口，让客户能够专注于业务本身，无需理解芯片更多底层细节就可实现模型的快速高效部署，MagicMind 插件化的设计还可以满足在性能或功能上追求差

异化竞争力的客户需求。

**七、公司此次软硬件协同迭代升级，提升了公司的产品竞争力，也体现了公司领先的核心技术水平，未来对于思元 370 乃至思元全系列产品的市场规划是什么？**

答：对于芯片公司，抢占市场、占据市场优势地位是至关重要的。短期来看，公司的主要目标是抓住战略机遇期，深耕行业客户，以目标客户需求为导向，将客户的开发需求有效融入公司的产品定义乃至技术创新，有针对性地推广和销售公司产品，为客户提供优质的产品和灵活高效的技术支持服务。从中长期角度来看，公司需要不断完善和优化已形成的云边端一体化、训练推理融合的基础系统软件平台和基于寒武纪产品的人工智能新生态，大力推进生态系统建设。公司希望基于统一的基础系统软件平台，彻底打破云端、边缘端、终端之间的开发壁垒，大幅提升人工智能应用在不同硬件平台的开发效率和部署速度，从而提升特定产品线的既有客户成为其他产品线的潜在客户的可能性，以期形成公司既有业务的坚固“护城河”。

目前，公司已经成为一家具备软硬件全栈系统能力的芯片设计公司，从云端推理思元 270、边缘推理思元 220、云端训练思元 290，到最新发布的推训一体思元 370，公司可为客户提供不同尺寸、全品类多场景的产品及统一的平台化基础系统软件，满足客户的差异化需求。其中，公司新产品思元 370 主要面向中高端推训场景，与主要面向训练的高端产品思元 290 形成协同，共同为客户提供全功能、全场景的智能算力。而且，思元 370 系列加速卡已与国内主流互联网厂商开展深入的应用适配，在语音、视觉等场景的性能表现较为出色。

未来，随着对客户支持的优化和完善以及软件生态的改善，公司新产品从导入到完成适配的周期将缩短。同时，公司与行业头部客户的良好合作将在市场上形成标杆或示范作用，

推动公司产品在更多标志性应用场景的落地。

**八、公司如何看待在智能驾驶领域的规划和布局，公司的竞争优势是什么？**

答：作为一家具备软硬件全栈系统能力的芯片设计公司，公司可充分利用在云边端智能芯片领域已有的技术积累，并基于统一的基础系统软件平台，调动云边端车的软件生态与应用场景联动，最大程度发挥车载智能芯片的性能，构建“云边端车”统一智能生态。

此外，随着智能驾驶应用场景的深入拓展，高等级智能驾驶必然产生更高的人工智能计算需求。智能驾驶作为新兴领域技术远未成熟，且高级别智能驾驶所需的基础硬件更处于发展早期阶段。目前公司子公司行歌科技尚在研发中的首款产品就是面向高等级智能驾驶应用场景的车载智能芯片。若未来行歌科技相关业务的开展情况良好，将进一步增强公司在智能芯片领域的技术积累和研发实力，成为公司目前主营业务的有益补充，推动公司的业务和生态拓展。

**九、近来越来越多互联网企业开始自研芯片，公司如何看待？是否会对公司业务产生一定影响？**

答：部分互联网企业研发的人工智能芯片主要是围绕其主营业务或针对特定应用场景完成完整或较为完整的行业解决方案，但是，在其他人工智能应用场景仍存在市场需求。而且，针对高端人工智能芯片领域，目前仅有少数公司推出较为成熟的智能芯片产品。公司已经成为一家具备软硬件全栈系统能力的芯片设计公司，可以为互联网公司提供不同尺寸、全品类多场景的产品及统一的平台化基础系统软件，满足客户的差异化需求。

**十、根据公司今年第三季度报告，从年初至报告期末的营业收**

	<p>入同比增长，主要系边缘智能芯片及加速卡销售收入增长所致。该类产品当前收入增长的主要原因是什么？另外，公司对产品线的未来规划如何？</p> <p>答：公司边缘智能芯片是面向边缘智能处理低延时、低功耗以及小尺寸要求的高性能芯片，可落地的应用场景广泛，支持边缘计算场景下的智能数据分析与建模、视觉、语音、自然语言处理等多样化的人工智能应用，其客户数量较多，因此在今年前三个季度处于持续推广及出货状态。</p> <p>未来，公司将持续推动云边端智能芯片产品线、基础系统软件平台的迭代升级与优化，并根据云边端各系列产品定位、市场发展前景与客户需求，有针对性地推广和销售公司产品，巩固已有客户、发掘潜在客户，持续构建和完善公司云边端一体化生态。</p> <p>十一、公司在人工智能芯片领域已经具备领先的核心技术，也有多款智能芯片产品落地，未来是否会保持研发投入？</p> <p>答：目前，人工智能芯片已受到多家集成电路龙头企业的重视，也成为多家初创集成电路设计公司的发力重点。随着越来越多的厂商推出人工智能芯片产品，该领域市场竞争日趋激烈。</p> <p>公司矢志成为业界领先的人工智能芯片设计公司。人工智能芯片属于计算芯片范畴，研发复杂的计算芯片需要持续充裕的研发投入，用于芯片底层基础技术的升级迭代、各产品线新产品的研发以及优秀人才的引进。公司将通过持续的研发投入提升产品的竞争力。</p>
附件清单（如有）	无
日期	2021 年 11 月 30 日