

第一节 重要提示

1.本年度报告摘要来自年度报告全文,为全面了解本公司的经营成果、财务状况及未来发展规划,投资者应当到http://www.sse.com.cn网站仔细阅读年度报告全文。

2.重大风险提示
公司已在本报告中描述可能存在的相关风险,敬请查阅本报告“第三节 管理层讨论与分析”之“四、风险因素”。

3.本公司董事会及董事、高级管理人员保证年度报告内容的真实性、准确性、完整性,不存在虚假记载、误导性陈述或重大遗漏,并承担个别和连带的法律责任。

4.公司全体董事出席董事会会议。
5.安永华明会计师事务所(特殊普通合伙)为本公司出具了标准无保留意见的审计报告。

6.公司上市时未盈利且尚未实现盈利
7.是否适用
8.董事会审议通过的本报告期利润分配预案或公积金转增股本预案
公司2025年度利润分配预案为:拟以实施权益分派股权登记日登记的总股本扣除公司回购专用账户上已回购股份后的股份余额为基数,每10股派发现金红利人民币3.90元(含税)。

公司2025年中期利润分配方案为:以实施权益分派股权登记日登记的总股本扣除公司回购专用账户上已回购股份后的股份余额为基数,每10股派发现金红利人民币2.00元(含税)。

公司2025年中期利润分配方案已于2025年10月实施完毕。此外,2025年度公司拟以现金方式,通过集中竞价方式回购股份金额为420,723,406.89元(不含佣金、过户费等交易费用)。

综上,本年度公司现金分红(含回购)的合计为1,119,349,610.08元(含税),占2025年度合并报表中归属于上市公司股东净利润的比例为60.7%。

母公司存在未弥补亏损
适用 不适用
8.是否存在公司治理特殊安排等重要事项
适用 不适用

第二节 公司基本情况

1.公司简介
1.1.公司概况
适用 不适用

Table with 4 columns: 控股股东, 实际控制人, 实际控制人, 实际控制人

1.2.公司存续凭证情况
适用 不适用

1.3.联系人和联系方式

Table with 4 columns: 姓名, 职位, 电话, 电子邮箱

2.报告期公司主要业务简介
2.1.主要业务、主要产品或服务情况

澜起科技是一家全球领先的无晶圆厂集成电路设计公司,致力于为客户提供AI基础设施提供创新、可靠及高效的互连解决方案。目前公司拥有两大产品线,互连类芯片和非互连类芯片。

Table with 2 columns: 互连类芯片, 非互连类芯片

互连类芯片产品线
1.内存互连芯片
(1)内存互连芯片(RCD/DB、MRCD/MDB及CXL)
内存互连芯片是服务器内存模组(又称“内存条”)的核心逻辑器件,作为服务器CPU存取内存数据的必经之路,其主要作用是提升内存存取访问的速度及稳定性,满足服务器CPU对内存模组日益增长的高性能及大容量需求。

DDR4及DDR5内存接口芯片按功能可分为两类:一是寄存时钟缓冲器(RCD),用来缓冲来自内存控制器的地址、命令、时钟、控制信号;二是数据缓冲器(DB),用来缓冲来自内存控制器或内存颗粒的数据信号。

DDR4及DDR5内存接口芯片按功能可分为两类:一是寄存时钟缓冲器(RCD),用来缓冲来自内存控制器的地址、命令、时钟、控制信号;二是数据缓冲器(DB),用来缓冲来自内存控制器或内存颗粒的数据信号。

澜起科技凭借自主知识产权的高速、低功耗技术,长期致力于新一代服务器平台提供符合JEDEC标准的高性能内存接口解决方案。随着JEDEC标准和内存技术的发展演变,我们先后推出了DDR2-DDR5系列内存接口芯片,可应用于各种规格内存模组,包括RDIMM及LRDIMM等,满足服务器对高速、大容量的内存存储的需求。

Table with 2 columns: RCD/DB芯片, MRCD/MDB芯片

图:RCD/DB芯片及含RCD/DB芯片的DDR5、LRDIMM内存模组示意图
我们的DDR4内存接口芯片子代产品及其主要应用情况如下:

Table with 2 columns: 主要应用, 主要应用

当前,DDR5内存模组已取代DDR4成为市场主流产品。DDR5是JEDEC标准定义的第五代双倍速率同步动态随机存取存储器标准,与DDR4相比,DDR5采用了更先进的工艺制程(11V),同时在传输有效性和可靠性上又迈进了一步。最新推出的DDR5第五代RCD芯片支持速率可达9000MT/s,是DDR4最高速率(3200MT/s)的2.825倍。

我们的DDR5内存接口芯片子代产品及其主要应用情况如下:

Table with 2 columns: 主要应用, 主要应用

DDR5第一代RCD芯片支持双通道内存架构,命令、地址、时钟和控制信号1:2缓冲,并提供奇偶校验功能。该芯片符合JEDEC标准,支持DDR5-4800速率,采用1.1V工作电压,更为节能。该款芯片除了可作为中央处理器专用用于RDIMM外,还可以与SDR5 DB芯片组成双通道,用于LRDIMM,以提供更兼容、更低功耗的内存解决方案。

DDR5第二代DB芯片是一款双向数据缓冲芯片,该芯片与DDR5 RCD芯片一起组成套件,用于DDR5 LRDIMM。该芯片符合JEDEC标准,支持DDR5-4800速率,采用1.1V工作电压,在DDR5 LRDIMM应用中,一颗DDR5 RCD芯片搭配十颗DDR5 DB芯片,即每个子通道配置五颗DB芯片,以支持片上数据缓冲,并将数据读取提升至最高16倍,从而为高速多核服务器提供更大容量、更高带宽和更佳性能的提升解决方案。

2022年9月,公司在业界率先发布DDR5第二代RCD芯片、DDR5第三代RCD芯片及双通道内存架构,命令、地址、时钟和控制信号1:2缓冲,并提供奇偶校验功能。该芯片符合JEDEC标准,支持数据速率56000MT/s,采用1.1V工作电压,更为节能。

2023年10月,公司在业界率先发布DDR5第三子代RCD芯片、DDR5第四子代RCD芯片支持的数据速率高达56000MT/s,较第二子代RCD速率提升14.3%,较第一子代RCD速率提升33.3%。

2024年1月,公司推出DDR5第四子代RCD芯片,DDR5第四子代RCD芯片支持的数据速率高达7200MT/s,较第三子代RCD速率提升12.5%,较第一子代RCD速率提升46%。

2024年第四季度,公司推出DDR5第五子代RCD芯片,DDR5第五子代RCD芯片支持的数据速率高达8000MT/s,较第四子代RCD速率提升11.1%,较第一子代RCD速率提升66.7%。

MRCD/MDB芯片
MRCD/MDB芯片是服务器高带宽内存模组MRDIMM的核心逻辑器件。随着AI及大数据应用的发展以及相关技术的演进,服务器CPU的内存数量快速增长,对内存储系统带宽的需求也日益迫切,以满足多核CPU中各个内存控制器的需求,MRDIMM正是基于这种应用需求而开发的。作为一种高带宽内存模组,第一代MRDIMM支持8800MT/s速率,第二代产品支持12800MT/s速率,每根MRDIMM模组均需搭配1颗MRCD、1颗SPD、2颗TS及1颗PMIC芯片。

MRDIMM的工作原理如下:MDB芯片用来缓冲来自内存控制器或DRAM内存颗粒的数据信号,在标准速率下,通过MDB芯片可以同时访问两个DRAM内存阵列(而传统RDIMM只能访问一个阵列),从而实现双倍带宽。MRCD则用来缓冲来自内存控制器的地址、命令、时钟、控制信号。MRDIMM的特点和优势包括:①使用专用缓冲的DRAM颗粒;②与现有DDR5生态系统有良好的适配性;③能够大幅提升内存模组的带宽。

我们的DDR5第一子代MRCD/MDB芯片于2024年开始在行业最广泛应用。我们的第二子代MRCD/MDB芯片已成功向全球主要厂商“送样”,该芯片支持12800MT/s传输速率,旨在为下一代计算平台提供卓越的内存储性能,满足云计算和人工智能等应用场景对内存带宽的迫切需求。

Table with 2 columns: MRCD/MDB芯片, MRDIMM内存模组

图:MRCD/MDB芯片及含MRCD/MDB芯片的MRDIMM内存模组示意图
我们的DDR5高带宽内存接口芯片(MRCD/MDB)及其主要应用情况如下:

Table with 2 columns: 主要应用, 主要应用

从下游应用来看,预计MRDIMM将在云计算、AI等内存带宽敏感的应用领域有较大需求。随着MRDIMM未来渗透率的提升,MRCD/MDB(尤其是MDB)芯片的需求也将大幅增长。

CXL芯片
在DDR4时代及DDR5代初期,时钟缓冲功能集成在RCD芯片中,用于服务器内存模组,尚未在PC端内存模组(如台式机和笔记本)部署。随着DDR5传输速率持续提升,时钟信号频率越来越高,信号完整性问题愈发显著。根据JEDEC定义,当DDR5数据速率达到9400MT/s及以上时,PC端内存模组需引入专用的时钟缓冲器(CKD,即“Clock Divider”)芯片,对时钟信号进行缓冲和重新驱动,以满足高速时钟信号的高性能和稳定性需求。

我们目前推出的第一代CKD芯片最高支持7200MT/s速率,旨在提高PC端内存模组的访问速度和稳定性,以匹配不断增长的CPU使用性能和需求。该芯片符合最新的JEDEC标准,支持双向总线地址和访问12、13C位。通过配置寄存器控制芯片,该芯片可将其输出特性切换成类似DDR4的网路拓扑,并可通过禁用未使用的信号引脚降低功耗。报告期内,我们推出新一代CKD芯片,最高支持9200MT/s速率,可有效优化客户端内存子系统性能,为下一代高性能PC、笔记本电脑及工作站提供关键技术支撑。

Table with 2 columns: CKD芯片, CXL内存模组

图:CKD芯片及含CKD芯片的CUDIMM内存模组示意图

澜起科技股份有限公司 2025 年度报告摘要

随着AI算力的迅猛增长,5G通信的持续升级及工业自动化的不断深化,市场对时钟信号的精度与稳定性要求日益严格。我们的时钟芯片产品凭借高可靠性、低功耗和广泛适配能力,已成功通过多家头部客户的测试验证,未来将广泛应用于AI服务器与数据中心、通信基础设施、工业控制设备、消费电子及汽车电子等领域。

公司代码:688008 公司简称:澜起科技

公司DDR5时钟缓冲器芯片及其主要应用情况如下:

Table with 2 columns: 主要应用, 主要应用

由于AI PC需要更高内存带宽以提升整体运算性能,AI PC渗透率的提升预计将加速DDR5的子代迭代,并推动对更高速率DDR5内存的需求。因此,AI PC应用的普及将加剧CKD芯片的需求提升。

根据JEDEC标准,DDR5内存模组除了内存颗粒及内存接口芯片外,还需要三种配套芯片,分别是串行检测集线器(SPD)、温度传感器(TS)以及数据管理芯片(PMIC)。

我们的内存模组配套芯片及其主要应用情况如下:

Table with 2 columns: 主要应用, 主要应用

● 串行检测集线器(SPD)
我们与合作伙伴共同研发了DDR5串行检测集线器(SPD),芯片内部集成了8Kbit EEPROM、12C/13C总线集线器(Hub)和温度传感器(TS),适用于DDR5系列内存模组(如RDIMM、RDIMM、MRDIMM、UDIMM、SODIMM、CUDIMM、CSODIMM、CAMM和LPCAMM等),应用范围包括服务器、台式机及笔记本内存模组。SPD是DDR5内存模组中不可或缺的组件,也是内存管理系统的关键组成部分,其包含如下几项功能:

第一,其内置BIOS EEPROM,是一个非易失性存储器,用于存储内存模组的相关信息以及模组上内存颗粒和相关器件的所有配置参数。根据JEDEC的内存规范,每个内存模组都需要配置一个SPD器件,并按其JEDEC规范的器件数据刷写SPD EEPROM的内容。主板BIOS在开机后会读取SPD内存存储的信息,并根据读取到的信息来配置内存控制器和内存模组。DDR5 SPD数据可通过12C/13C总线访问,并按可擦写区域(block)进行保护,以确保DDR5内存模组的稳定性和数据安全。

第二,该芯片还可作为12C/13C总线集线器,一端连接系统主控设备(如CPU或基板管理控制器(BMC)),另一端连接内存模组上的本地组件,包括RCD、PMIC和TS,是系统主控设备与内存模组上本地组件之间的通信中心。在DDR5规范中,一个12C/13C总线最多可连接两个集线器(8个内存阵列),每个集线器与该集线器管理的每个内存模组上的本地组件都被绑定了一个特定的地址代码,支持唯一地址固定寻址。

第三,该芯片还内置了温度传感器(TS),可连续监测SPD所在位置的温度。系统设备可通过12C/13C总线从SPD中的相关寄存器读取温度传感器检测到的温度,以便于进行内存模组的温度管理,提高系统工作的稳定性。

● 温度传感器(TS)
我们与合作伙伴共同研发了DDR5高精度温度传感器(TS)芯片,该芯片符合JEDEC规范,支持2C和3C串行总线,适用于DDR5服务器内存模组(如RDIMM、LRDIMM和MRDIMM)。TS作为SPD芯片的从设备,可以工作在时钟频率分别高达1MHz、12C和12.5MHz、13C总线上;CPU可通过SPD芯片与之进行通讯,从而实现对内存模组的温度管理。TS是DDR5服务器内存模组上重要组件,通常一条内存模组配置两颗。

● 电源管理芯片(PMIC)
我们的DDR5电源管理芯片(PMIC)符合JEDEC规范,其主要作用是作为内存模组上的其他芯片(如DRAM、RCD、SPDRITs等)提供电源管理,CPU可通过SPD芯片与之进行通讯,从而实现电源管理。我们的DDR5 SPD、TS、PMIC芯片在内存模组中的位置示意图如下:



2. PCIe/CXL互连芯片
(1)PCIe Retimer芯片
PCIe Retimer芯片是适用于PCIe高速数据传输的超高速时序校正合芯片。近年来,随着PCIe协议从3.0(8 GT/s)发展到4.0(16 GT/s)、5.0(32 GT/s),并逐步迈向6.0(64 GT/s)和7.0(128 GT/s),数据传输速率的不断提升带来了显著的信号衰减和时钟抖动等问题,这些问题极大地限制了PCIe协议在下一代计算平台中的应用范围,促使行业加大对高速电路与系统互连技术的研究,同时也推动了超高速传输链路下信号完整性研发工作。

为了补偿信号传输的损耗,提升信号质量,超高速时序校正合芯片(Retimer)应用而生。目前,PCIe Retimer芯片已成为高速传输链路中不可或缺的重要组件,尤其在数据中心的数据高速、远距离传输场景中,可有效解决信号时序不匹配、串扰严重、完整性差等问题。

PCIe Retimer芯片采用先进的信号调理技术,能够补偿信号损耗并消除各种抖动源的影响,从而提升信号完整性,增加高速信号的有效传输距离。为服务器、存储设备及硬件加速器等应用场景可扩展的高性能PCIe互连解决方案。我们推出的PCIe Retimer产品组合包括PCIe 4.0 Retimer及PCIe 5.0/CXL 2.0 Retimer,其中,PCIe 4.0 Retimer芯片符合PCIe 4.0基本规范,支持16 GT/s的传输速率,可补偿高达28 dB的信道损耗;PCIe 5.0/CXL 2.0 Retimer符合PCIe 5.0和CXL 2.0基本规范,支持32 GT/s的传输速率,可补偿高达36 dB的信道损耗,支持业界主流封装,其功耗、传输延迟等关键性能指标达到国际先进水平,并已与CPU、PCIe交换芯片、固态硬盘、GPU及网卡等进行了广泛的互操作测试。报告期内,我们推出了PCIe 6.x/CXL 3.x Retimer芯片,支持64 GT/s的传输速率,采用PAM4 SerDes技术,支持高达43 dB的信道损耗,可用于通用及AI服务器、有源线缆(AE)和存储系统等典型应用场景。

我们的PCIe Retimer芯片及其对应板卡的位置示意图如下:



图:澜起科技数据保护和可信计算加速芯片
澜起科技数据保护和可信计算加速芯片采用公司自主创新的Monk-TSSSE可信安全系统扩展架构和技术,将硬件级数据加解密和可信度量两大核心功能融合于单一芯片之上。芯片内部集成了高速加解密、安全SoC和硬件信任(HiROT)三个子系统。

该芯片支持硬件SM2/3/4、SHA-256/384/12、AES、RSA、ECC等商用算法和国际主流加解密算法,还可广泛应用于对数据保密性、完整性要求极高的场景,如AI训练和推理、分布式数据存储、零信任安全等。芯片内置多个真随机数发生器(TRNG),搭配PCIe 5.0 x 8高速接口,可提供高达1600bps的吞吐量,多颗芯片集成可实现数据吞吐量倍增,从而为数据中心提供高性能加解密支持,助力服务器算力在数据中心中落地应用。

该芯片广泛兼容多项可信计算标准,具备出色能效及可信优势。芯片支持TPM、TCM和ITPCM等可信计算标准,遵从从IBM® 0008-2012、GMPT 0012-2020、GMPT 0028-2014等多项设计、测试、接口标准,并兼容PIPS-140设计规范和RINIST、SPM 800-193固件安全标准,支持MCTP/SPDM等协议。芯片可作为硬件信任根(HiROT)使用,满足可信平台3.0规范的需求,保障服务器、台式机、嵌入式终端、加速卡(AI卡)等各类设备启动运行期间的平台安全。

该芯片可用于解决数据中心等高并发数据加解密需求的需求,同时因其具有低功耗特点,也适用于边缘、边缘侧、嵌入式或嵌入式中等及低功耗安全需求的市场。

澜起科技数据保护和可信计算加速芯片
澜起科技数据保护和可信计算加速芯片采用公司自主创新的Monk-TSSSE可信安全系统扩展架构和技术,将硬件级数据加解密和可信度量两大核心功能融合于单一芯片之上。芯片内部集成了高速加解密、安全SoC和硬件信任(HiROT)三个子系统。

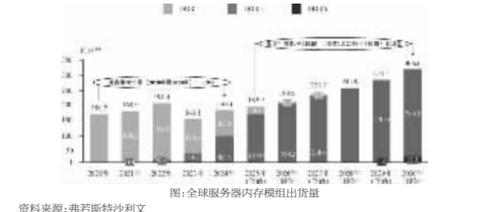


图:全球服务器内存模组出货量

资料来源:弗若斯特沙利文
(3)内存互连芯片行业情况
内存互连芯片包括内存接口芯片、内存模组配套芯片等。内存接口芯片是服务器内存模组的核心逻辑器件,其主要作用是提升内存存取访问的速度及稳定性,满足服务器CPU对内存模组日益增长的高性能及大容量需求。

内存接口芯片的发展演变情况如下:

Table with 2 columns: 主要应用, 主要应用

从2016年开始,DDR4技术进入成熟期并成为内存市场的主流技术。为了实现更高的传输速率和支持更大的内存容量,JEDEC进一步完善了DDR内存接口芯片的技术规格,增加了多种功能以支持更高速率和更大容量的内存。在DDR4时代一共有四个子代产品,每一子代内存接口芯片的最高传输速率不断提升,其中最后一个子代产品支持的最高传输速率达3200MT/s。随着DDR5内存技术的成熟商用,DDR5已取代DDR4成为市场主流产品。相比DDR4最后一个子代产品,DDR5内存接口芯片采用了更先进的工艺制程(11V),并在传输有效性和可靠性上进一步提升。根据JEDEC公布的信息,DDR5内存接口芯片已经经历了六个子代,前五个子代产品支持速率分别是4800MT/s、5600MT/s、6400MT/s、7200MT/s、8000MT/s,最后一个子代产品支持速率达9200MT/s。通过持续的技术创新,以实现更高的传输速率和支持更大的内存容量,将是内存接口芯片行业未来发展的趋势和动力。

在DDR5时代,根据JEDEC定义,服务器内存模组除了需要内存接口芯片之外,还需要配置三种配套芯片:一颗SPD芯片、一颗PMIC芯片和一颗TS芯片;普通台式机及笔记本电脑的内存模组(UDIMM、SODIMM)则需要配置两种配套芯片:一颗SPD芯片和一颗PMIC芯片。

随着技术的发展,内存互连领域衍生出新的接口芯片类型,包括用于服务器的高带宽内存模组MRDIMM的MRCD/MDB芯片,以及用于PC端内存模组的CKD芯片。

在服务器端,随着人工智能和大数据处理等应用快速发展,处理器内核数量日益增多,对内存带宽的需求急剧增长。JEDEC制定了新型高带宽内存模组多路复用双列直插内存模组MRDIMM(Multiplexed Rank DIMM)的相关技术规范。根据JEDEC公布的信息,DDR5 MRDIMM通过新设计提高了数据传输速率和整体性能。多路复用允许将多个数据信号组合并通过单个通道传输,从而在不增加外部物理接口的情况下提升带宽,实现无缝带宽升级。按数据速率超过同代的DDR5 RDIMM。其特点包括:①平台与RDIMM兼容,提供灵活的部署配置;②采用标准的DDR5 DIMM封装(包括DRAM、外引线、引脚、PMIC和TS),便于推广;③利用CDIO逻辑处理管理方面的功能(如扩编、自帮助刷写的LRDIMM)系统管理设计和技术挑战。MRDIMM未来将持续迭代升级,第一代MRDIMM支持8800MT/s速率,第二子代产品支持12800MT/s速率,正在定义的第三代MRDIMM支持的速率将可实现16000MT/s。MRDIMM需要搭配1颗MRCD和10颗MDB芯片,其设计复杂度要求高于普通的双列直插DDR5芯片。

在PC端,随着DDR5传输速率持续提升,到DDR5时代,原本无需信号缓冲的UDIMM、SODIMM(主要用于台式机和笔记本芯片),将需要配套一颗CKD芯片,对内存模组的时钟信号进行缓冲和重新驱动,从而提高时钟信号的完整性和可靠性。JEDEC已制定了CUDIMM和CSODIMM内存模组相关标准,包括CKD芯片标准,将应用于支持9400MT/s及以上内存模组的台式机和笔记本电脑。

Table with 2 columns: CKD芯片, CXL内存模组

图:全球内存互连芯片市场规模

资料来源:弗若斯特沙利文
(4)PCIe互连行业情况
PCIe协议是一种高速串行计算机扩展总线标准,自2003年诞生以来,其互连技术在近几年发展迅猛,传输速率基本3~4年翻倍增长,并保持良好的向后兼容性。从PCIe 4.0到PCIe 5.0,传输速率从16GT/s提升至32GT/s;PCIe 6.0到PCIe 7.0,传输速率将进一步提升至64GT/s和128GT/s。凭借强大的生态系统,平台厂商、芯片厂商、终端设备厂商和测试设备厂商深度合作,PCIe已成为主流互连接口,全面覆盖了PC机、服务器、存储设备等各种计算平台,广泛应用于云计算、企业级计算、人工智能和物联网等应用场景。

PCIe Retimer芯片
PCIe Retimer芯片是在PCIe协议升级迭代背景下应运而生,它主要解决数据中心和服务器在通过PCIe协议进行高速、远距离传输时,面临的信号时序不匹配、失真大、完整性差等问题。随着技术的发展,PCI协议协议迭代更新,传输速率不断提升,但信号完整性管理尺寸受限于工业标准,变化不大。这种导致整个链路的数据传输速率下降,传输速率下降,完整性下降,进而影响整个系统的性能。因此,业界亟需解决PCIe信号完整性的问题,以提高信号传输速率。

一种解决方案是选用数据缓冲PCB,但成本高昂且难以有效提高多连接器应用场景;另一种解决方案是引入数据缓冲芯片,如PCIe Retimer芯片,通过提升信号完整性管理技术和重新驱动技术,Retimer芯片能够补偿信号损耗并消除抖动影响,从而提升PCIe信号的完整性,增加高速信号的有效传输距离,相较于其他替代方案,Retimer芯片在性能、标准化和生态系统等方面具有明显优势,可用于PCIe与高速外设(如GPU、AI芯片、SSD卡及网卡等)的互连,适应多连接器应用场景,未来还可根据标准规范灵活切换至PCIe或CXL模式,更受用户青睐。

因此,随着传输速率提升,PCIe 4.0到16GT/s翻倍至PCIe 5.0的32GT/s,Retimer芯片的技术优势愈发显著。根据行业发展趋势,到CXL 3.0时代,PCIe Retimer芯片已成为行业主流解决方案。近两年,随着AI服务器需求快速增长,PCIe Retimer芯片已成为AI服务器中的关键部件,其市场规模也随之迅速扩大。

PCIe Switch芯片
PCIe Switch芯片是一种用于扩展连接多个PCIe设备的关键组件,可以将有限的PCIe通道分配给更多设备,同时优化带宽分配。

PCIe Switch芯片应用于实现高速、低延迟的设备互连的关键组件,其主要功能为:①扩展接口:可增加PCIe Express端口数量,让更多设备接入PCIe总线网络通信。如服务器中,当CPU的PCIe通道不足时,PCIe Switch芯片可提供更多设备,例如SSD、网卡、GPU等;②数据转发:在点对点(P2P)工作模式下,为连接的多设备提供数据传输通道,将多个设备的数据转发到目标设备上,实现设备高速连接;③实现冗余功能:当与AI或Switch连接的链路出现故障(VLAN),可将多余流量转发到另一片PCIe Switch,并进行负载均衡,把流量均匀分配给多个服务器,避免单一管理和服务节点,避免多个操作设备共用同一根PCIe总线的角色由不同网络地址冲突,支持NTBB(Non-transparent Bridge)技术,通过本地翻译实现不同网内设备之间的设备通信。例如,传统服务器系统中的多个控制器,可利用NTBB通过PCIe链路直接通信,实现数据和控制信息的互通。

PCIe Switch芯片可以突破传统PCIe接口的制约,实现更多高性能PCIe互连,显著提升了带宽和资源配置能力,在数据中心、云计算、存储设备、网络设备中有广泛的应用,尤其适用于对带宽和延迟敏感的场景。比如在数据中心和云计算中,PCIe Switch芯片可以连接更多PCIe Switch芯片进行互连,也可连接GPU/AI加速卡/网络接口卡进行互连;在存储系统中,PCIe Switch芯片可以连接大量NVMe SSD,构建超高速存储网络;在服务器端,PCIe Switch芯片可以连接多100G/400G网卡,管理更复杂的网络转发。

AI服务器的高速增长显著拉动了PCIe Retimer芯片和PCIe Switch芯片的需求。以配置两块GPU的典型AI服务器为例,通常需配备2至4个PCIe Switch接口扩展,同时需配备8至16个Retimer芯片,以延长CPU与外设间的有效传输距离。目前,根据部分国内头部CPU服务器厂商的配置,需要24个Retimer芯片。因此,PCIe互连芯片已成为AI服务器中不可或缺的核心器件,其需求与AI服务器出货量息息相关。

根据弗若斯特沙利文的数据,PCIe互连芯片市场规模从2022年的4.69亿美元增长将达到2024年的22.82亿美元。行业预测未来几年该市场规模将持续高速增长,预计2030年市场规模将达到77.61亿美元,2025至2030年间年均复合增长率高达20.1%。2024年,中国占全球市场25%以上的份额,预计到2030年将占30%以上。

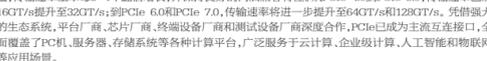


图:全球PCIe互连芯片市场规模

资料来源:弗若斯特沙利文
(5)CXL互连行业情况
作为一种新兴的高速互连技术,CXL自推出以来备受业界关注。随着人工智能、云计算等领域的高速发展,内存扩展、内存池化等CXL技术的应用正在受到越来越多厂商的积极部署,以打破内存瓶颈。

近年来,CXL技术在数据中心和人工智能领域展现出巨大的应用潜力。在数据中心领域,CXL技术通过高带宽、低延迟的特性,将不同的计算和存储资源进行互连,形成庞大的资源池,可以显著提升数据处理的效率,满足实时数据分析和处理的需求。在人工智能领域,CXL技术通过支持GPU和FPGA等加速器与主机处理器的有效协同,可显著提升AI模型训练和推理的速度,实现低延迟、高速的数据传输,从而大幅提升AI推理效率;同时CXL技术支持内存扩展和内存共享,为AI应用提供更大的内存空间和更灵活的内存访问方式。根据相关服务器厂商调研,CXL内存池化方案在AI推理、向量数据库和内存数据湖三个最重要的内存应用场景中,均具有卓越性表现;CXL内存池化方案为运行更大参数AI模型提供了更高性能和稳定的内存支持。

据CXL协会2024年发布的CXL3.2标准,2025年又继续推出了CXL4.0标准,CXL技术的相关生态也在不断完善:(1)从主流CPU厂商来看,英特尔发布了两款支持CXL 2.0协议的CPU(Granite Rapids和Sierra Forest),AMD发布了支持CXL 2.0协议的最新五代EPYC处理器,上述CPU平台将支持更多的CXL设备类型,提供更好的安全性和可靠性,适配更多的应用场景;2025年主流CPU厂商也将继续推出CXL3.0新平台的产品,未来几年CXL技术将在服务器平台中持续进一步提升。(1)从内存模组厂商来看,英特尔和三星电子正在积极研发并量产CXL兼容的内存模组,如三星电子展示了CMM-D 2.0模组,可将内存与处理器一起工作在128GB CMM-D内存模组内,而且计划在2026年推出CMM-D 3.0模组。(II)从服务器平台来看,CXL Switch芯片开始应用于服务器平台,支持PCIe协议的数据传输,可以实现一个节点内的内存和芯片设备资源的共享,未来,越来越多的服务器CPU和GPU都将支持CXL接口,这一趋势将显著推动CXL市场的发育,尤其是在数据中心、人工智能和云计算等领域的应用。

根据弗若斯特沙利文的数据,2024年CXL互连芯片市场规模处于商业化初期,市场规模约为430万美元,行业预测未来几年该市场规模将持续高速增长,预计至2030年市场规模将达到170.92亿美元,2025至2030年间年均复合增长率高达170.2%。2024年,中国占全球市场25%以上的份额,预计到2030年将占30%以上。

