



链滴

# CCIX 标准简介

作者: [xiaoye45](#)

原文链接: <https://ld246.com/article/1680565269884>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

\* [<https://zhuanlan.zhihu.com/p/375519357> - 知乎专栏](<https://zhuanlan.zhihu.com/p/375519357>)

\* 随着众核时代的到来，系统架构由SoC演进到Chiplet。Chiplet的实现分为物理层、接口层、通信和协议层，具体在正文中有描述，这里着重介绍一下协议层中的CCIX标准。随着机器学习和大数据应用的高速发展，仅以中央...

\* 2023-01-06 09:16:12

随着众核时代的到来，系统架构由SoC演进到Chiplet。Chiplet的实现分为物理层、接口层、通信层和协议层，具体在正文中有描述，这里着重介绍一下协议层中的CCIX标准。

随着机器学习和大数据应用的高速发展，仅以中央处理器为中心的服务器架构无法满足这些应用的性能需求，因此需要高效的异构计算架构，如图形处理单元 (GPU)、可编程逻辑阵列 (FPGA)、智能网卡 (NIC) 和很多其它领域特定的可编程器件等。对目前很多计算任务，加速器能够比单独的处理器度更快、功耗更低的完成所需功能。但是，不受管控的异构会带来软件复杂性。当前主要采用PCIe作处理器和片外加速器间传输数据的协议，但面对基于不同指令集(ISA)的处理器或应用特定的加速器如提升带宽、降低时延、增加易用性，成为下一代互联的首要诉求。

为解决下一代互联问题，由处理器、加速器和其它外设厂商促成了CCIX联盟的成立，希望通过变革加速器的内存和数据移动的方式制定出新的互联标准，来保证片间无缝传输，实现片外加速器互联。

适用于加速器的缓存一致性互联标准 (CCIX, Cache Coherent Interconnect for Accelerators, 也成 "see 6") 采用两种机制来提高性能、降低延时。第一种是采用缓存一致性，自动保持处理器和加速器的缓存一致，提升易用性、降低延时；第二种是提高CCIX链接的原始带宽，最高链接速率升至25 T/s(千兆传输/秒)。CCIX规范也规定了多个CCIX端口如何聚合，以提供超过单个接口的性能，匹配加速器和内存扩展带宽。

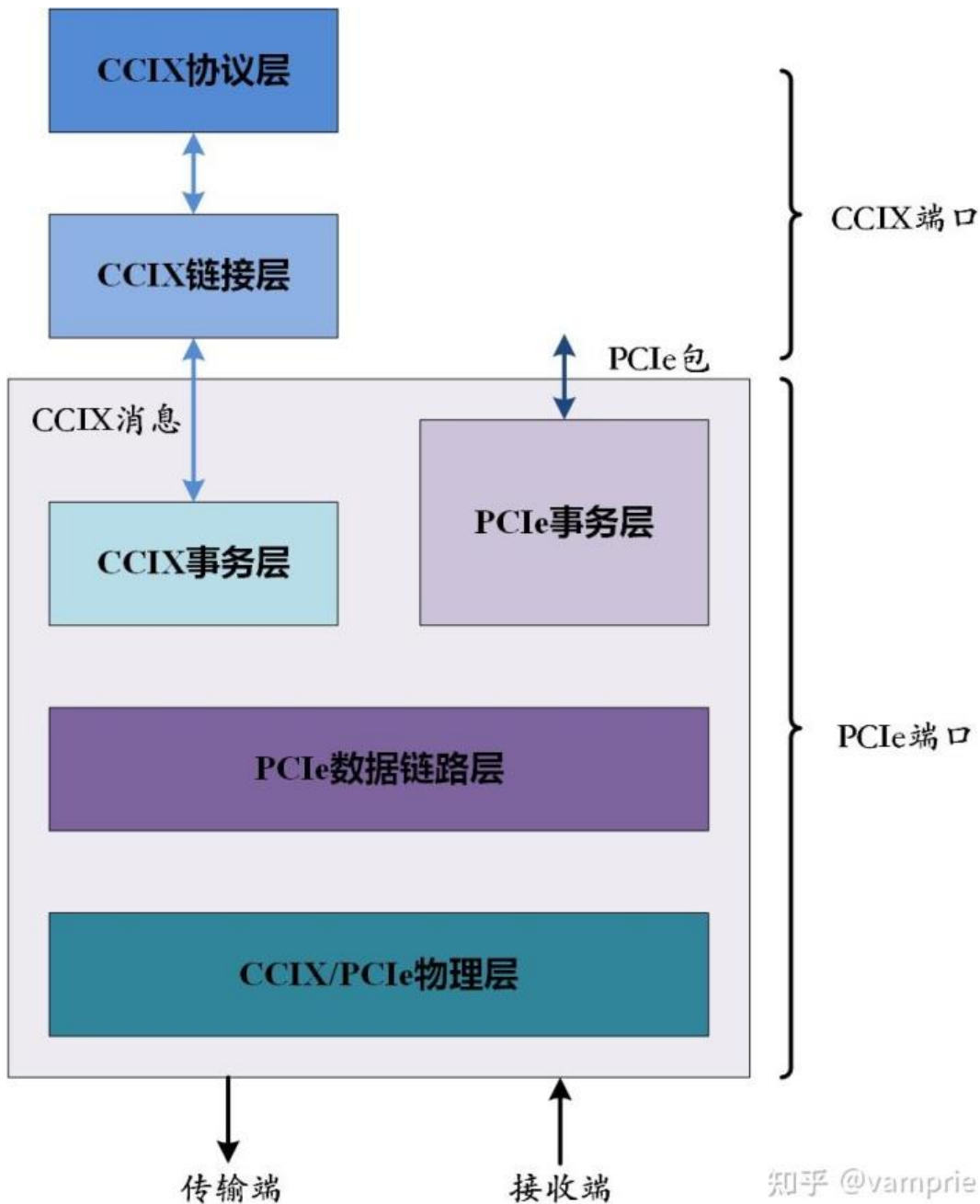
在多核系统中，针对缓存一致性问题已经有了很多技术手段 (如ARM的CCI中采用的监听过滤器等) CCIX将其基本原理扩展到加速器，应用数据就可以在处理器缓存和加速器缓存间自主传递，不需要软驱动参与。除了缓存，CCIX还支持将PCIe 设备自带的内存扩展为系统内存。CCIX的数据共享模型是于虚拟地址 (VA) 寻址的共享内存。

处理器和加速器的缓存和/或内存可通过CCIX协议自动更新和同步，因此只需要传递数据指针而不需行数据搬移。从而减小数据延时，提升应用性能。同时减轻软件负担，使其聚焦于应用处理而不是底数据传输。

## 1. CCIX架构

CCIX基于PCIe架构扩展而来，采用分层架构，如图9所示，主要包含协议规范和传输规范。其中，协规范包含CCIX协议层和CCIX链接层，规定了缓存一致性协议、报文发送、流控和传输部分的协议。输规范包含CCIX和PCIe事务层、CCIX PCIe数据链路层和CCIX物理层，负责器件间的物理连接，包速率和带宽协商、传输包错误检测和重试和初始包编码协议。

□



知乎 @varnprie

□

图9 CCIX分层架构

**\*\* CCIX协议层\*\***

负责一致性协议，包括内存的读、写等，也提供片上一致性协议（如Arm的CHI）的简单映射。通过定义的缓存状态使硬件能够确定当前内存的状态，如数据是否唯一且未被修改（和内存一致），或共且被修改（和内存不一致）等。

**\*\* CCIX链接层\*\***

负责CCIX协议层定义的代理（agent）之间消息的传输格式。目前构建在PCIe之上，采用分层架构，来可映射到不同的传输层。此外，还负责端口聚合，使得多个端口能够聚合在一起提升带宽。

**\*\* CCIX和PCIe事务层\*\***

分别负责处理来自CCIX和PCIe的包。

PCIe协议支持部署虚拟通道，使得不同数据流通过一个PCIe链路。将CCIX和PCIe传输流各分到一个拟通道，使得CCIX和PCIe传输可以共享相同的链路。

CCIX能够传输标准的PCIe包或经过优化的CCIX包。经过优化的CCIX包删减了PCIe包里的几个不必要字段。传输标准的PCIe包时可以采用现有的PCIe交换器。传输经过优化的CCIX包，能降低PCIe的额外开销，使得一致性传输的包更小、更高效。

## \*\* PCIe数据链路层\*\*

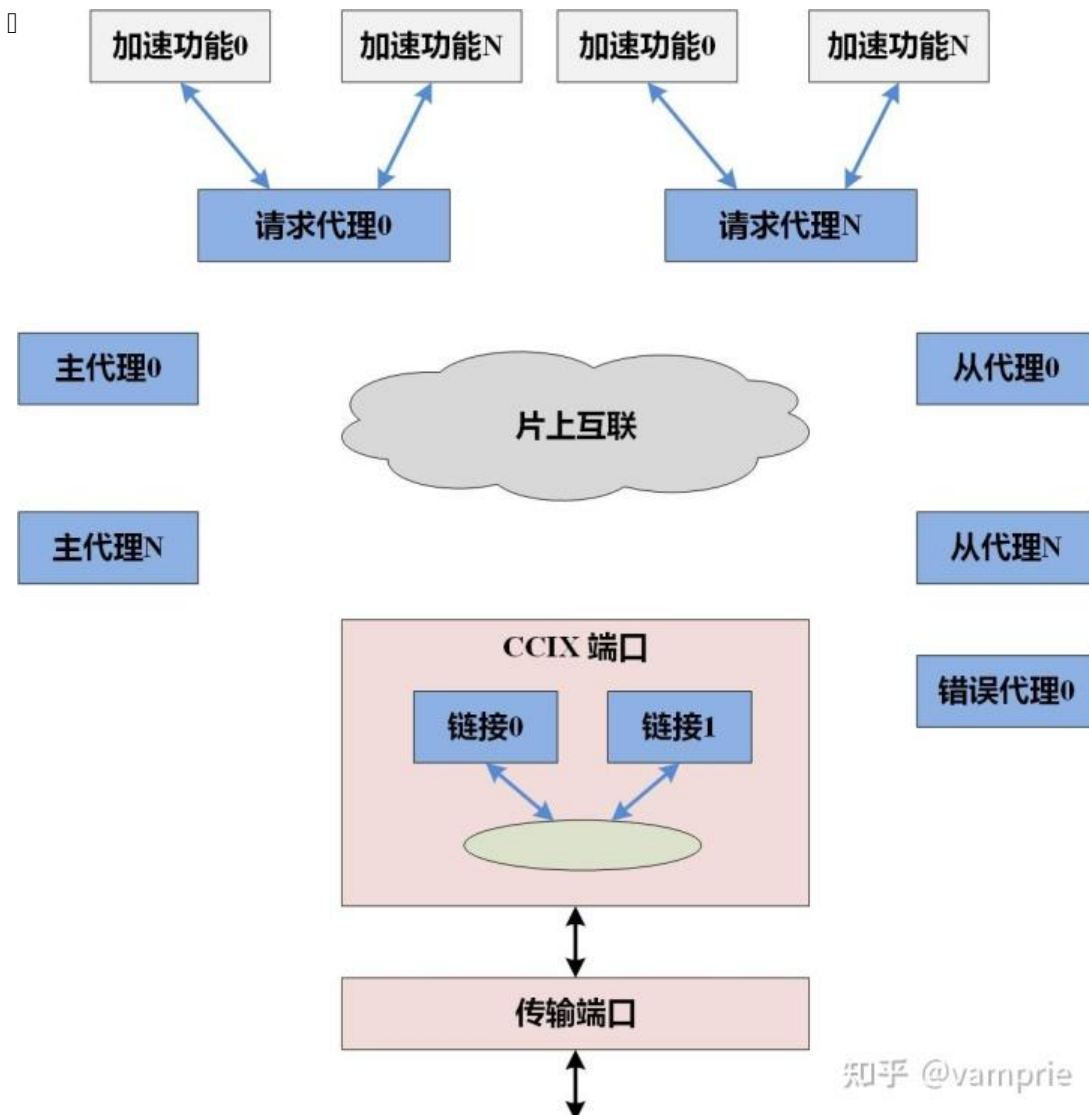
执行数据链路层的所有正常功能，包括CRC错误校验、包确认和超时检查和信用初始化及交换。

## \*\* CCIX/PCIe物理层\*\*

基于PCIe物理层进行扩展，速率可达到25GT/S（千兆传输/秒）。该模式称为ESM（扩展速率模式）会在两个具备ESM模式的器件连接时自动识别，比PCIe4.0的速率（16GT/S）提升56%。

## 2. CCIX组件

CCIX协议定义了CCIX组成模块的内存访问协议。所有CCIX器件至少有一个具备CCIX链接的端口。一个CCIX端口关联一组物理管脚，用于和另一个CCIX端口连接，在两个或多个不同芯片间交互信息。包的组件如图10所示，有请求代理（RA）、主代理（HA）、从代理（SA）和错误代理（Error Agent），协议中为每种代理都定义了唯一的代理ID值。



## 图10 CCIX架构组件

### \*\* 请求代理 (RA) \*\*

请求代理可以对系统内的不同地址进行读、写操作，也可以对已经访问过的数据进行缓存。每个请求代理的发起源可以是一个或多个处理单元。请求代理为加速器或CCIX使能的IO主设备提供了一致性系统内存接口，使得加速器的缓存具备一致性，此过程对软件透明。

### \*\* 主代理 (HA) \*\*

主代理负责管理指定的一段地址的数据一致性。当一个缓存行的状态需要改变时，主代理通过向相关请求代理发出侦听操作来保持一致性。

### \*\* 从代理 (SA) \*\*

CCIX支持将外设所附内存扩展为系统内存。当主代理在一个芯片上，而与其有关的内存在另一个芯片时需要进行内存扩展，扩展内存所需的架构组件为从代理。从代理不会被请求代理直接访问，而是需通过一个主代理进行中转。

### \*\* 错误代理 (EA) \*\*

一个错误代理接收并处理CCIX组件发出的协议错误信息。

## 3. CCIX特点

得益于CCIX的分层架构，能够支持多种灵活的拓扑结构。最常见的拓扑结构是直接连接的共享虚拟内存。其它拓扑结构，如星型、交叉或网格拓扑，也很容易被构建和支持；

由于具备缓存一致性，支持主设备和加速器间的数据共享和内存扩展，可以无需驱动实现数据搬移，低了延时和软件开销。同时这种共享数据结构，可以减少对不同加速器的单独控制和管理，实现统一任务调度，简化软件程序，提高协同运行效率。

## 小结

CCIX联盟是一个开放的工业标准组织，制定独立于任何处理器架构的、非私有的CCIX标准。CCIX标为内存扩展器件和加速器提供高性能、低延时、芯片与芯片间的互联。特别为应对未来数据中心、云算、大数据及其它需要异构计算应用的巨大挑战而设计。CCIX标准将是革命性的进步，它拓展了开放异构架构和缓存一致性共享内存模型的优势，来满足未来数据中心不断演进的需求。