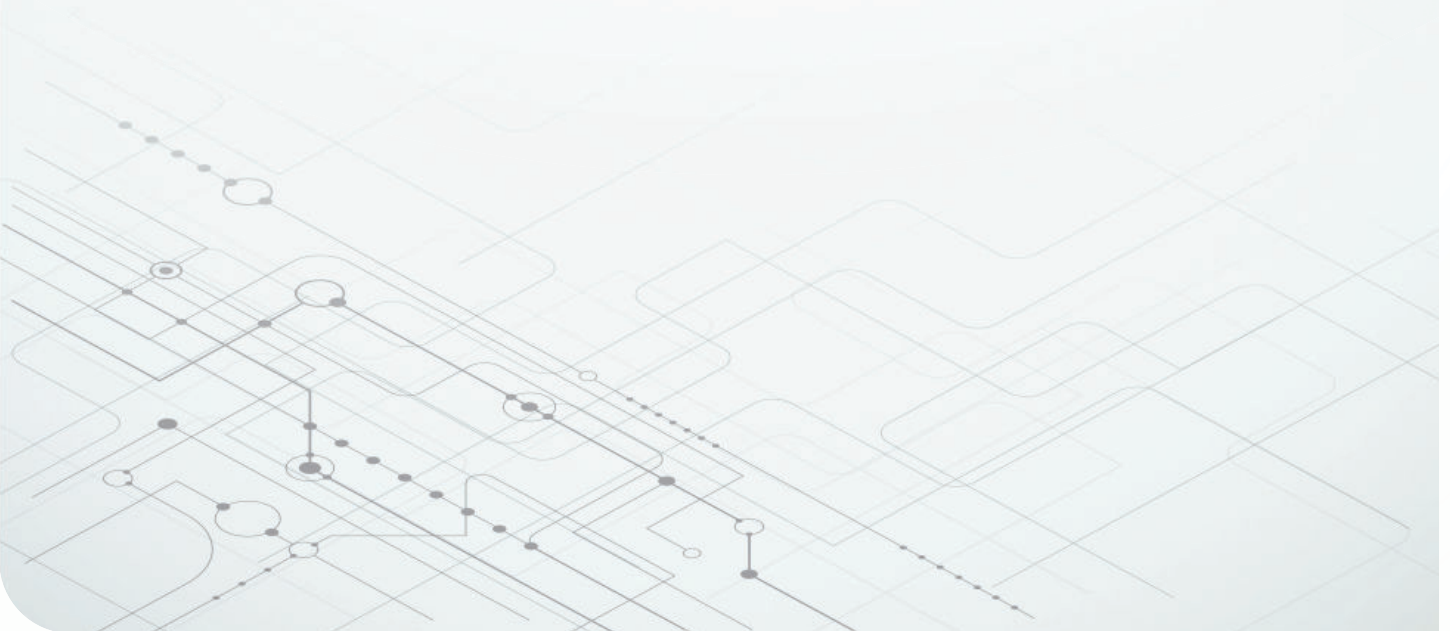


时钟在数据中心的的重要性



Hao Zheng
Systems Engineer



内容概览

- 1 深入了解数据中心**
数据中心通过服务器和交换机处理、存储和中继数据。
- 2 数据中心的时钟**
快速外设组件互连 (PCIe) 时钟架构有两种类型，在数据中心内有多种用途：通用时钟 (CC) 和独立参考时钟 (IR)。
- 3 更低抖动是大势所趋**
随着以太网数据速率的增加，高速串行器/解串器需要更低的抖动。
- 4 更高的集成度**
BAW 集成可提高可靠性、降低抖动并缩减尺寸和成本。

人工智能 (AI) 和各种云服务正在推动数据中心的发展。AI 培训需要更多计算资源，而媒体流等云服务需要更多的存储和数据处理。

存储和计算的增长需要更高的连接速度。因此，PCIe 6.0 链路数据速率提高到 64GT/s，以太网通道速度高达 224Gbps。速度更高的数据链路需要噪声更低的时钟来保持高数据质量。

深入了解数据中心

如图 1 所示，数据中心通常由多个机架的服务器组成。每个服务器机架的顶部都有一个 ToR 交换机，用于在服务器和网络之间中继数据包。主干或结构交换机是将 ToR 交换机连接到网络的高层交换机。

高数据速率通常需要在服务器与 ToR 交换机之间使用有源电缆，并在 ToR 交换机与主干或光纤交换机之间使用光学模块来降低损耗。图 2 展示了通常需要计时组件的服务器块。

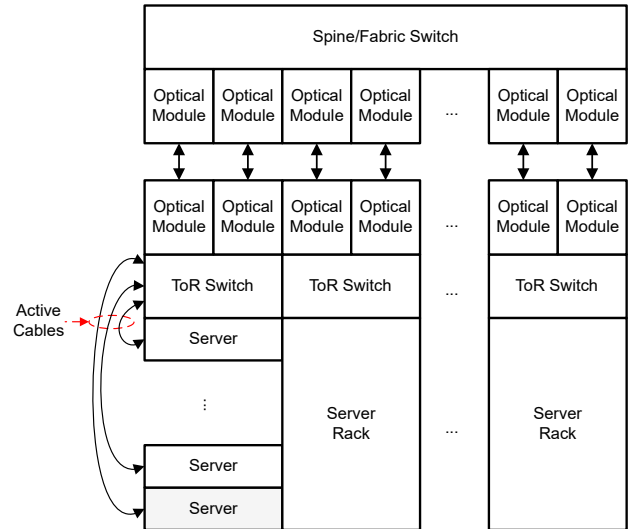


图 1. 数据中心架构。

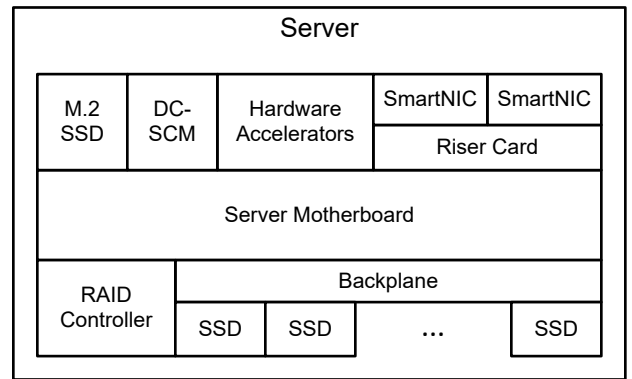


图 2. 服务器块。

数据中心的时钟

图 3 和图 4 分别展示了服务器主板上的 PCIe 内部时钟和外部时钟架构。在内部时钟服务器上，CPU 生成 PCIe 时钟。然后，这些 PCIe 时钟由 PCIe 时钟缓冲器扇出。缓冲器输出为各种端点计时，或通过 PCIe 连接器传递到子卡。

外部时钟服务器上可能有多种 PCIe 时钟源：来自外部时钟发生器的本地 PCIe 时钟或由 CPU 生成的 PCIe 时钟。每个端点或连接器都可以从这些源中选择一个，具体取决于它们所属的时钟域。

器件和接口通常需要由振荡器或时钟发生器提供的低频单端时钟。

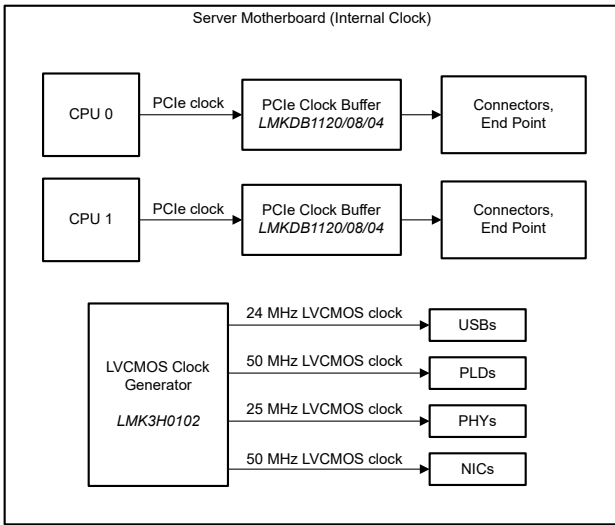


图 3. 服务器主板内部时钟架构示例。

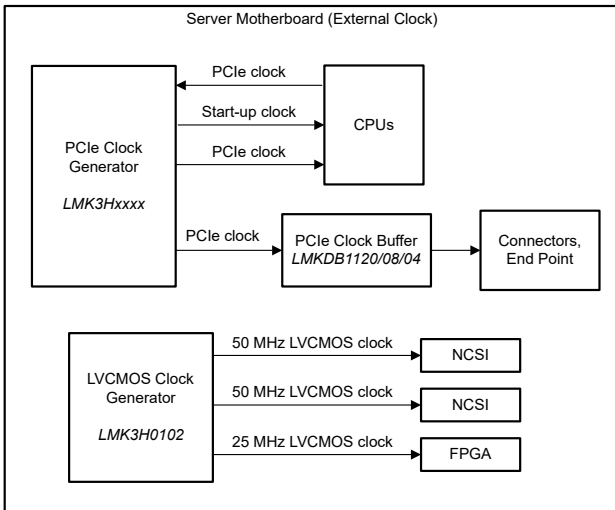


图 4. 服务器主板外部时钟架构示例。

数据中心安全控制模块 (DC-SCM) 是由 Open Compute Project 定义的附加卡。在图 5 所示的示例中，从服务器主板为 DC-SCM 卡提供了 PCIe 参考时钟。然而，基板管理控制器和 USB 主机控制器都需要 PCIe 时钟。您不能简单地将布线分开并将一个时钟路由到两个器件，因为这会使振幅减半并降低信号完整性。因此，时钟信号将不再满足 PCIe 合规性要求，这就是需要 PCIe 时钟缓冲器的原因。时钟缓冲器会接收一个时钟输入并生成多个输入副本，而不会降低信号完整性。

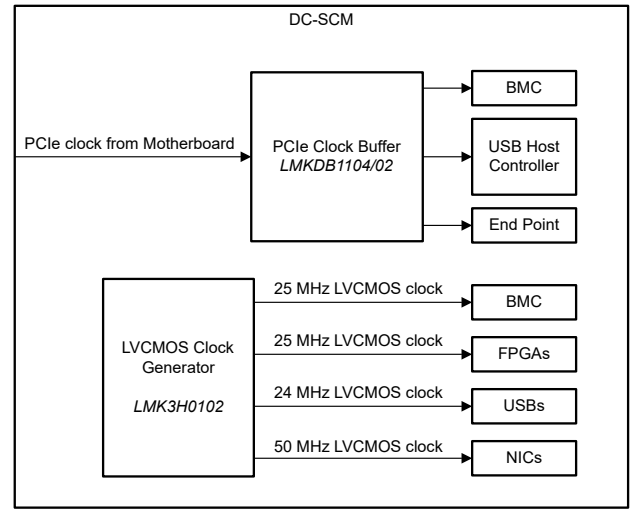


图 5. DC-SCM 时钟架构示例。

与 DC-SCM 类似，其他扩展卡或 PCIe 插入卡也可能需要时钟缓冲器来分配 PCIe 时钟，如图 6 所示。

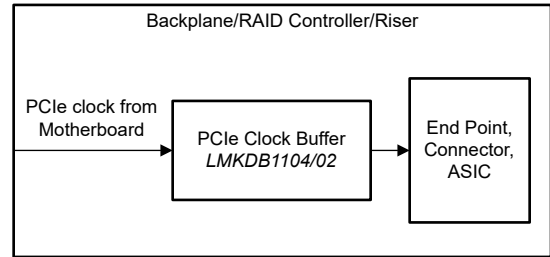


图 6. PCIe 插入卡时钟架构示例。

网络接口卡 (NIC) 将服务器连接至网络。SmartNIC 提供了额外的计算资源来减轻服务器 CPU 的负载。NIC 和 SmartNIC 都需要 PCIe 和以太网时钟。在图 7 所示的示例中，有两个 PCIe 时钟源：一个是来自 CC 架构主板的通用 PCIe 时钟，另一个是用于 IR 架构的本地 PCIe 时钟。在正常工作模式下，NIC 在 CC 上运行。但是，如果 CC 丢失或不可用，NIC 可以切换到 IR 并改为使用本地 PCIe 时钟。此外，由于 NIC 通过以太网端口连接到交换机，应用特定集成电路中的以太网串行器/解串器通常需要高性能的 156.25MHz 时钟。

用于 AI 培训等特定计算任务的硬件加速器的 PCIe 时钟要求与 SmartNIC 类似。图 8 展示了一个 PCIe 时钟架构示例。PCIe 时钟仅在本地生成，而不是采用 CC 和 IR 架构并在两者之间进行切换。在该示例中，CPU、图形处理单元和其他端点需要许多时钟。因此，一个双通道时钟发生

器以及两个 20 通道时钟缓冲器可以生成多达 40 个 PCIe 时钟。硬件加速器不需要以太网时钟，因为它未连接到 SmartNIC 等 ToR 交换机。可能存在除 PCIe 以外的专有链路，该链路可能需要额外的高性能时钟，但与以太网时钟类似。

图 9 是仅使用 IR PCIe 架构的另一个示例。双通道 PCIe 时钟发生器用于为固态硬盘 (SSD) 控制器计时。

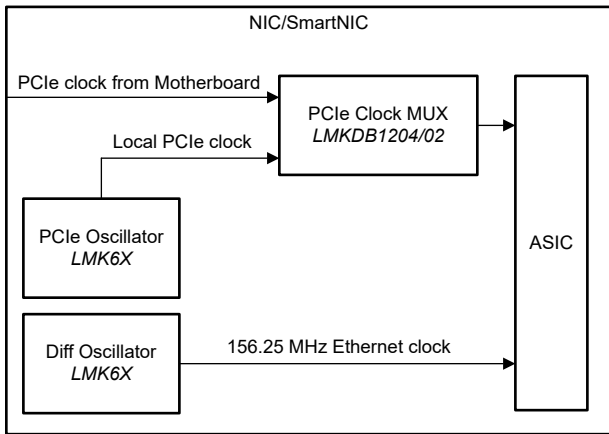


图 7. NIC 和 SmartNIC 时钟架构示例。

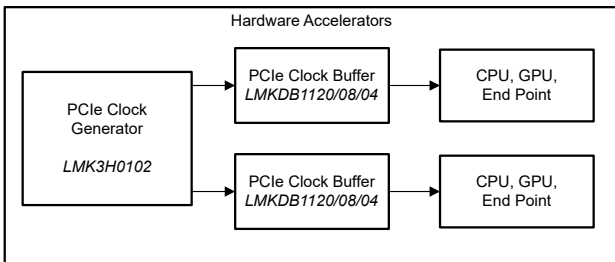


图 8. 硬件加速器时钟架构示例。

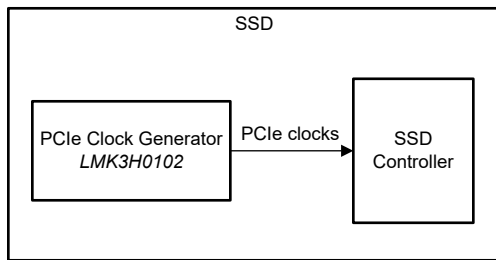


图 9. PCIe SSD 时钟架构示例。

以太网通道速度为 56Gbps 或更高会显著影响任何无源电缆的插入损耗。因此，需要“主动”互连来降低损耗和提高数据质量。根据距离的不同，有不同类型的主动互连。有源电缆（包括基于铜的有源电缆和基于光纤的有源光

缆）可以短距离连接，例如在 NIC 与 ToR 交换机之间连接。

光学模块用于较长距离的连接。还有不同类型的光学模块。其中一些用于数据中心内的 ToR 交换机和主干或光纤交换机之间，而另一些则可用于数据中心之间。

由于以太网通道速度较高，光学模块数字信号处理器需要一个极低噪声的以太网时钟，如图 10 所示。另一方面，有源电缆中的以太网重定时器只需要一个常规时钟，如图 11 所示。

1PPS 信号携带时钟同步信息，并从主干或光纤交换机向下传递到 ToR 交换机，然后传递到 NIC 或 SmartNIC。您可能需要在有源电缆焊盘中使用时钟缓冲区或电平转换器进行电平转换并生成其他副本。

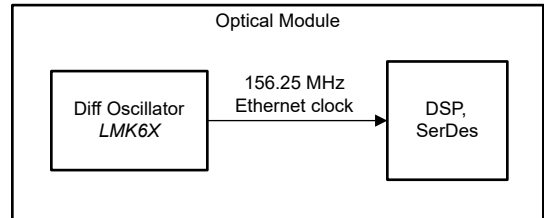


图 10. 光学模块时钟架构示例。

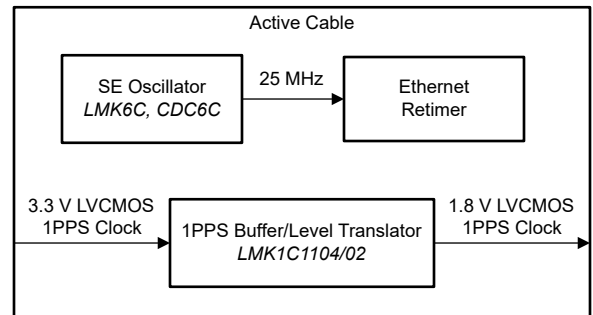


图 11. 有源电缆时钟架构示例。

为高速以太网串行器/解串器生成参考时钟需要一个抖动极低的时钟发生器，如图 12 所示。主干或光纤交换机还需要与 ToR 交换机类似或更好的以太网时钟性能。此外，您还需要一个用于网络同步的计时数字锁相环 (DPLL)，如图 13 所示。

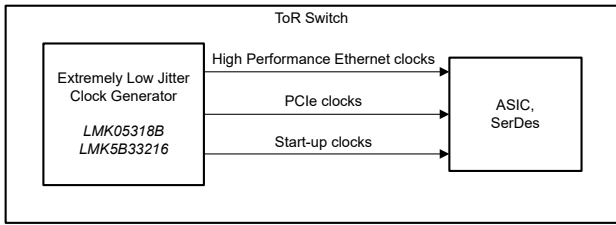


图 12. 简化的 ToR 交换机时钟架构。

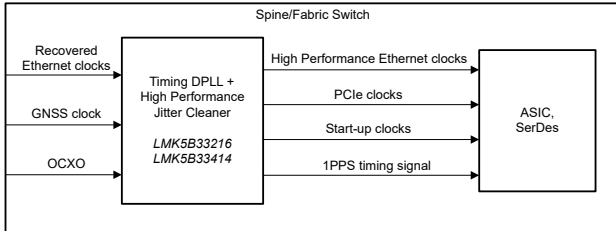


图 13. 简化的主干或光纤交换机时钟架构。

更低抖动是大势所趋

对于高速以太网应用（例如 ToR 交换机、主干或光纤交换机、光学模块以及 NIC 或 SmartNIC）来说，更低的抖动是必不可少的。112Gbps 通道速度以太网串行器/解串器通常需要频率为 156.25MHz 时具有 125fs 或 100fs 12kHz 至 20MHz 的均方根抖动。224Gbps 以太网通常需要 70fs；未来的以太网时钟应实现最大值为 50fs 的抖动或更佳的抖动。TI 专有的体声波 (BAW) 技术在当前的 TI 产品中可在 156.25MHz 频率下提供最大值为 65fs 的抖动。参考资料 [1] 提供了有关 BAW 技术的更多细节。

PCIe 参考时钟的抖动要求也越来越严格，尤其是在 PCIe 第 6 代中，其中调制方案从非归零 (NRZ) 更改为脉冲振幅调制 4 级 (PAM-4)。与 NRZ 中的两个电平相比，PAM-4 在四个电平上运行，因此它需要降低基准时钟的噪声。也正是由于此原因，56Gbps PAM-4 以太网需要明显优于 28Gbps NRZ 的抖动性能。但是，由于 PCIe 合规性定义了噪声传递函数来“滤除”抖动，因此 PCIe 的抖动要求比以太网宽松得多。

更高的集成度

数据中心的另一个趋势是时钟集成，它可以实现更高的可靠性、更小的尺寸和更低的成本。在空间有限的子卡上，缩小面积尤其重要。

集成时的一个重要步骤是消除外部晶体谐振器 (XTAL) 或晶体振荡器 (XO)。一些供应商提供集成了 XTAL 的 IC，但这种集成有一些缺点。首先，晶体集成通常需要具有基板的特殊 Land Grid Array (LGA) 封装，而不是更简单的业界优选 Quad Flat No-Lead (QFN) 封装。LGA 封装成本更高，而且不利于焊料检查。此外，将晶体堆叠在基础芯片顶部会增加封装高度。例如，常规 QFN 封装的高度仅为 0.9mm。由于集成了晶体，封装高度可达 1.7mm，这对于外壳而言可能是个问题。

集成 BAW 可以避免所有这些问题。BAW 不仅提供低于 65fs 的 RMS 抖动，而且体积非常小，成本也很低。将 BAW 芯片放在基板芯片上不需要 LGA 封装；常规的 0.8mm 或 0.9mm QFN 封装高度就足够了。因此，BAW 集成的成本远低于晶体集成。此外，与晶体相比，BAW 对振动的敏感性较低，老化性能更好，并且故障率也低得多。数据和详细信息可在参考资料 [2] 和 [3] 中找到。

集成还使设计人员能够组合缓冲器、多路复用器和本地时钟。单个集成电路可以生成本地 PCIe 时钟、缓冲或多路复用外部 PCIe 时钟，同时对 1PPS 信号进行电平转换。一个时钟发生器即可满足您的所有时钟需求，而无需使用多个时钟发生器、多路复用器、缓冲器和振荡器。

结论

数据中心时钟树正变得越来越复杂，需要各种时钟器件，包括振荡器、缓冲器、多路复用器、时钟发生器和网络同步器。一些应用需要低成本和低性能的时钟，而另一些应用则要求很少有供应商能够实现的极高性能。

TI 旨在提供可满足任何数据中心时钟要求的完整时钟产品系列，通过更高的集成度简化时钟树，通过更低的抖动提高系统性能，并利用 BAW 技术降低物料清单 (BOM) 成本。

参考资料

1. 德州仪器 (TI): [TI BAW 技术可在高速网络中实现超低抖动时钟](#)
2. 德州仪器 (TI): [高可靠性 BAW 振荡器 MTBF 和时基故障率计算](#)
3. 德州仪器 (TI): [独立 BAW 振荡器相对于石英晶体振荡器的优势](#)
4. [时钟缓冲器](#)
5. [时钟发生器](#)
6. [时钟抖动清除器和同步器](#)
7. [振荡器](#)

重要声明: 本文所提及德州仪器 (TI) 及其子公司的产品和服务均依照 TI 标准销售条款和条件进行销售。建议客户在订购之前获取有关 TI 产品和服务的最新和完整信息。TI 对应用帮助、客户的应用或产品设计、软件性能或侵犯专利不负任何责任。有关任何其它公司产品或服务的发布信息均不构成 TI 因此对其的认可、保证或授权。

所有商标均为其各自所有者的财产。

重要声明和免责声明

TI“按原样”提供技术和可靠性数据（包括数据表）、设计资源（包括参考设计）、应用或其他设计建议、网络工具、安全信息和其他资源，不保证没有瑕疵且不做任何明示或暗示的担保，包括但不限于对适销性、某特定用途方面的适用性或不侵犯任何第三方知识产权的暗示担保。

这些资源可供使用 TI 产品进行设计的熟练开发人员使用。您将自行承担以下全部责任：(1) 针对您的应用选择合适的 TI 产品，(2) 设计、验证并测试您的应用，(3) 确保您的应用满足相应标准以及任何其他功能安全、信息安全、监管或其他要求。

这些资源如有变更，恕不另行通知。TI 授权您仅可将这些资源用于研发本资源所述的 TI 产品的应用。严禁对这些资源进行其他复制或展示。您无权使用任何其他 TI 知识产权或任何第三方知识产权。您应全额赔偿因在这些资源的使用中对 TI 及其代表造成的任何索赔、损害、成本、损失和债务，TI 对此概不负责。

TI 提供的产品受 [TI 的销售条款](#) 或 [ti.com](#) 上其他适用条款/TI 产品随附的其他适用条款的约束。TI 提供这些资源并不会扩展或以其他方式更改 TI 针对 TI 产品发布的适用的担保或担保免责声明。

TI 反对并拒绝您可能提出的任何其他或不同的条款。

邮寄地址：Texas Instruments, Post Office Box 655303, Dallas, Texas 75265

Copyright © 2023，德州仪器 (TI) 公司