1. **背景** [1]

NVLink是英伟达（NVIDIA）公司开发的用于近距离半导体通信的基于线的通信协议，可用于cpu和gpu之间的处理器系统和gpu之间的数据和控制代码传输。NVLink指定一个点对点连接，每个方向的数据通道的数据速率分别为20和25 Gbit/s (v1.0/v2.0)。实际系统中的总数据速率为160和300 GByte/s (v1.0/v2.0)，表示输入和输出数据流的系统总数。到目前为止，NVLink产品主要介绍了高性能的应用程序空间。NVLINK于2014年3月首次发布，使用Nvidia开发的专有高速信令互连(NVHS)。

1. **总线部分特性** [1]

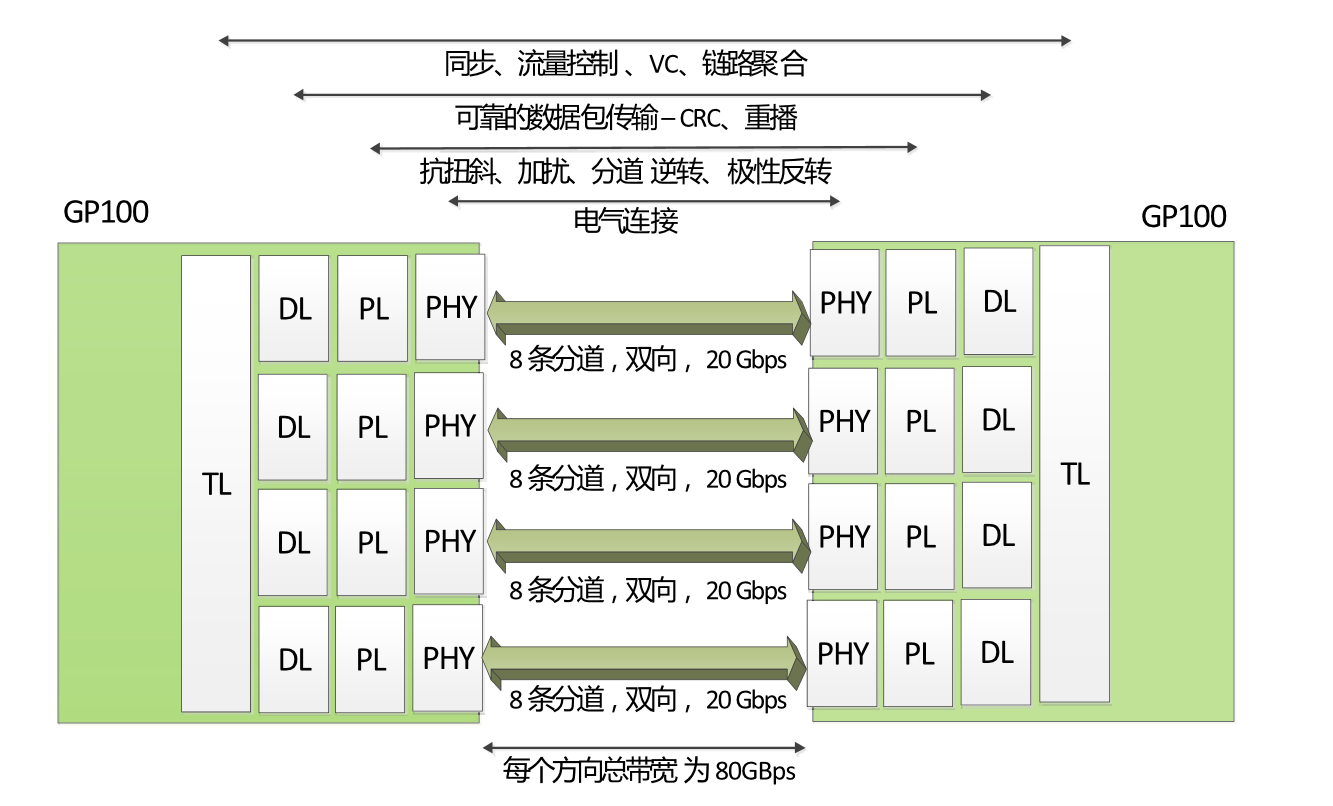
尺寸非常小，约为PCIe的1/3。

每个NVLink包括八个不同方向的差动对，总共32根电线。

1. **NVLink信号与协议技术** [2]

**NVLink控制器层**

NVLink控制器由3层组成，即物理层（PL），数据链路层（DL）以及交易层（TL）。该协议使用可变长度数据包，数据包大小范围从1（如简单的读请求指令）到18个（用于包含地址扩展的256位数据传输的数据写请求）流量控制位。下图展示了NVLink的隔层和链路；物理层（PHY）、数据链路层（DL）、交易层（TL）。



**物理层（PL）**

物理层与物理级电路连接。

作用：抗扭斜（所有8条分道）、制定框架（找出每个数据包的起始点）、加扰/解扰（确保足够的数为转移密度以支持时钟恢复）、极性反转、分道逆转以及向数据链层传输接收到的数据。

**数据链路层（DL）**

作用：在整个链路上可靠地传输数据包。以及负责链路初启和维护；可将数据继续发送至交易层。

带传输的数据包利用25位CRC（循环冗余校验）进行保护。已传输的数据包保存在重播缓冲区中，直到这些数据包被链路的另一端的接收方积极认可（ACK）为止。如果数据链路层在接收的数据包上检测到CRC错误，则不发送ACK并准备接收重新发送的数据。数据包只有在得到认可后才会从重播缓冲区退出。25位CRC最多可以检测5个随机数位错误或者在任何分道上最多检测25位突发错误。CRC是在当前的数据头以及之前的有效负载（如果有的话）上进行计算的。

**交易层（TL）**

作用：处理同步、链路流量控制、虚拟通道并且可以聚合多条链路以便在处理器之间提供极高的通信宽带。

Source：

[1] 维基百科URL：https://en.wikipedia.org/wiki/NVLink

[2] NVIDIA Tesla P100白皮书，附录A，p35-p36