

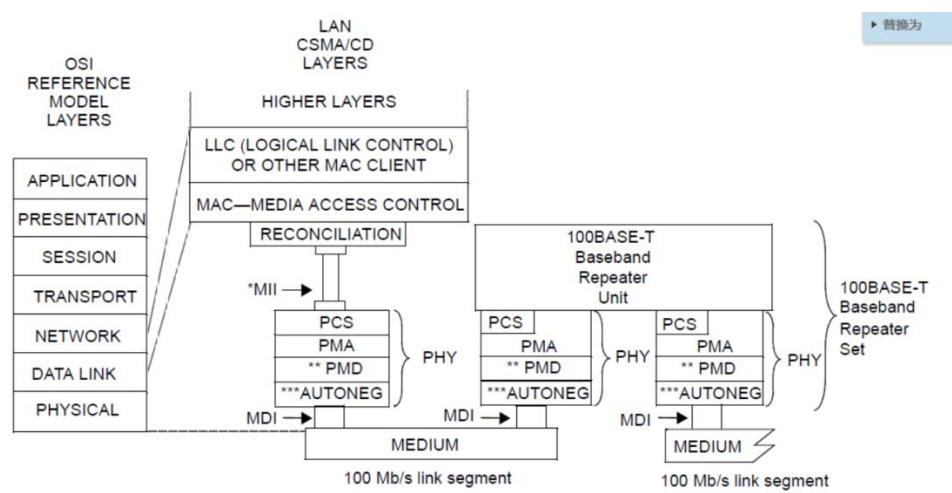
以太网的物理层实现

以太网的媒体访问控制器，物理层编码芯片，介质和 MDI 接口、信号编码、MII 接口

引言

为了能适应现代信息化社会长距离、高速率和高吞吐量的通讯需求，以太网协议设计了专门的传输介质和传输编码。作为一个讨论嵌入式产品如何实现网络连接的系列科普文章（兼本人的学习笔记），本章以国际电子电气工程师协会发布的 IEEE802.3 为基准依据，以以太网具体使用的现状事实为切入点开展叙述。在 IEEE802.3 系列协议中，以太网物理层被进一步划分，如图 1-2 所示：

图 1-2 以太网物理层的结构设计



Linkspeed 不同的以太网物理层结构也是不同的。以百兆为例，可以看出，物理层被划分为 PCS(物理编码子层)、PMA（物理媒体访问）、PMD（物理介质关联层接口，这个可选）、AUTONEG（自动协商层，这个也是可选的）这几个子层，我觉得 MDI（介质相关接口）和 MEDIUM（介质）也可以认为是 OSI 七层模型中物理层的一部分。上图的右半部分应该是个集线器之类的东西。其中 PCS、PMA、PMD 和 AUTONEG 功通组成一个完整的物理层芯片，可以看出物理层芯片就是通过 MII 接口接 MAC 层，通过 MDI 接口连接介质这样完成接收波形，解码，校验，编码成 MII 时序再转发这样一个工作流程。

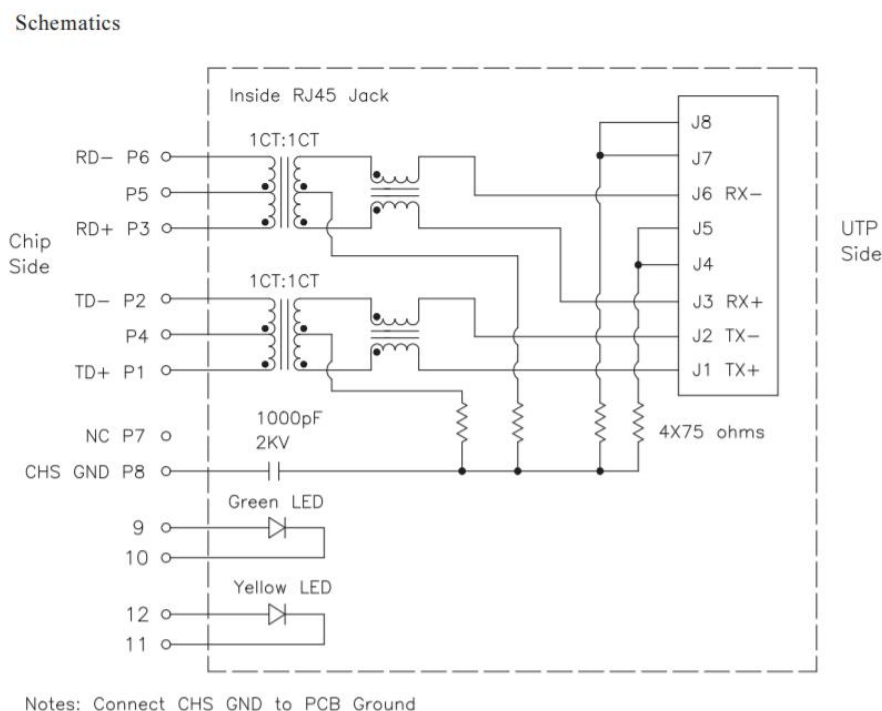
物理层芯片和 MDI 接口

在搭建嵌入式系统的电路设计中,当今市场上有一系列产品来实现以太网,已经非常便捷。在这里我将嵌入式设备实现以太网连接分为这么几个组件或设备:媒体访问控制器(即 MAC)、物理信号编码芯片(即 PHY,或称为物理层芯片)、网络变压器,网口,双绞线四个部分。本章我们将讨论日常用到的网线,网口连接器,网络变压器,物理信号编码和实现物理信号编码的芯片,不讨论除数据帧之外的各种流控帧,也不讨论半双工模式,而媒体访问控制器将放到下章数据链路层来讨论。

本章只讨论 UTP 双绞线(即常说的电口连接的铜质缆线)做以太网介质的情况,不讨论光纤。在实际使用中,10M/100M 以太网使用的是两对差分线,而 1000M 则需要四对差分线。在五类和六类网线中,剥开最表面的绝缘皮,你可以发现有 4 对双绞线,线色为橙绿蓝棕。经常可以听闻“网线只要接一半就可以用”的言论,其实这只是局限在 link 速率在 10M 和 100M 的情况下,而目前家庭铺设局域网都是千兆起步万兆为佳的,你书房墙上的网线面板还只接四根线就跑不到千兆了。而千兆的连接速率需要用满网线的四对共八根双绞线。注意,这里的 10M/100M/1000M 连接速率指的是位(比特)每秒,即 bps,换算成字节还需要除以八。

在网线工作在 10M/100M 的连接速率时,有一对双绞线专门负责发送的,另一对专门负责接收。这种能同时进行收发的通讯方式是标准的全双工模式,不过如果物理层芯片进行速度协商时,协商成了半双工模式,那么就只有一对双绞线被使用了。嵌入式产品中最常用的网络变压器和 RJ45 网口连接器合一的产品,汉仁(HanRun)的 HR911105A 的结构图体现了这点。

图 1-3 HR911105A 的结构图



看结构图的右侧可以看出，连接插入的五类线或者六类线的那八根弹簧件，即右侧连接 J1 至 J8 的部分，只有第 1，2，3，6 这四根线实际连接了网络变压器（即图中两个 1:1 的那两个变压器），剩下的第 4，5，7，8 这四根线直接连接了滤波电路，实际上没有起到传输信号的作用，HR911105A 是只能用在 10M 和 100M（专指最常用的 100BASE-TX 标准）的连接速率下的。再看左侧的 P1，P2，P3，P6 这个接线 PAD，就是连接物理层芯片的，可以看到他们分别标注了 TD+,TD-,RD+,RD-的名称，也就设计用来发送数据的差分线的 positive 和 negative 端，和接收数据的差分线的 positive 和 negative 端。但是这种需要区分发送和接收两个方向的情况现在已经改变了。

由于这种接收、发送分别独立成两组线的分线设计方式，在制作网线时，需要像串口一样将两端的收和发相对而接。计算机与计算机，计算机与交换机的连接，有时候需要用交叉网线，有时候需要用直连网线。但人们实际使用中，并没有关注是要用交叉线还是直连线，大都是随便买来直接接在各个设备的网口上了。

为什么如此还是能工作正常呢？

因为现在的以太网物理层芯片支持一个功能，称为“MDI/MDIX 自动翻转”，即物理层芯片的接收端和发送端可以自动对调，所以即使网线的接收发送两个方向接反了也没事。

但千兆以太网则不是这种情况，千兆以太网以 UTP 双绞线做介质时是要用满 4 对双绞线的，而且每一对都同时用作接收和发送（专指最常用的 1000BASE-T 标准），因此不能以 10M/100M 的眼光看待千兆以太网的网线，网口和物理层芯片。

上文提到的“MDI”是介质相关接口的意思，两个以太网物理层芯片之间的接口就是 MDI 接口，双绞线，网络变压器和物理层的 MDI 侧都是在这个范畴。当把网线插到网口上之后，也就是物理层建立电气连接之后，物理层芯片会进行速度协商，然后会确立一个网线两端的物理层芯片和网线都能接受的速度，当然了网线质量越差协商出的速度自然越慢。但一般就五种速度，10M 全双工和半双工，100M 全双工和半双工，还有 1000M（全双工）。

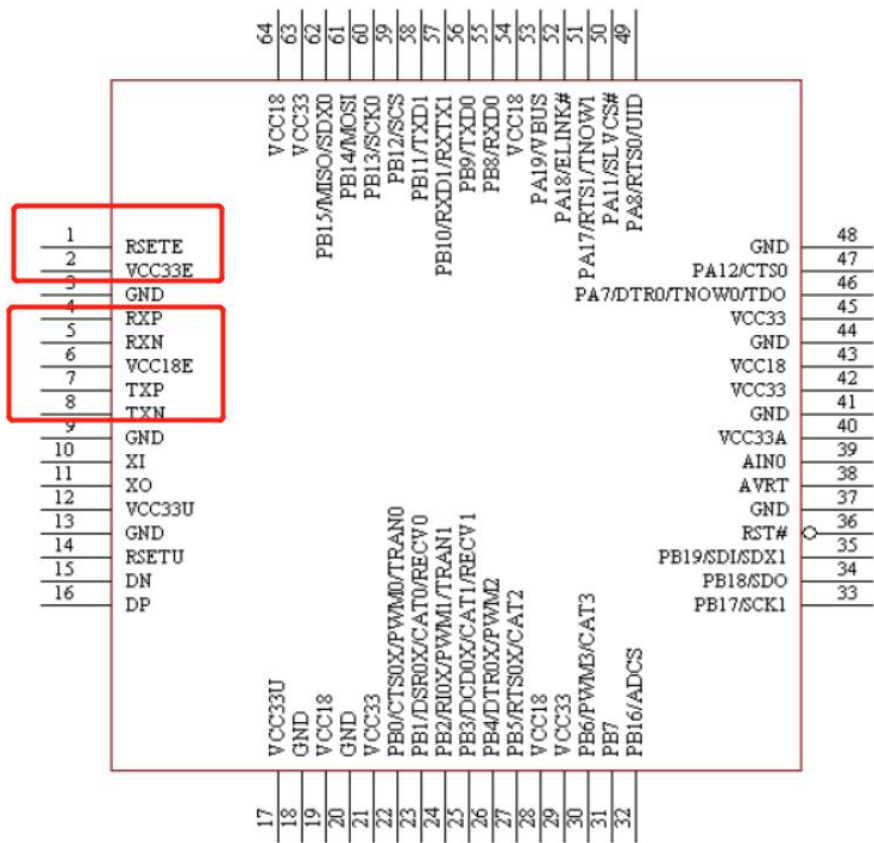
MDI 接口上的电信号已经是由物理层编码好的信号，可以直接发送发到介质上去了。这个介质一般指双绞线，具体由物理层的功能决定。很多物理层支持通过光模块把光纤当作介质，光纤的通信性能比双绞线更好，且价格更低，但是光模块的价格一直居高不下。

一般人们在 MDI 接口从物理层芯片引出后，连接双绞线之前之前，会加网络变压器进行隔离。从图 1-3 中的结构图中可以看出，HR911105A 自带两对绕组网络变压器。网络变压器的组成可以看成是一个初级线圈和次级线圈匝数一致的隔离变压器串联一个共模电感，而其存在的意义则是将设备电路和介质进行隔离和滤波。

市场上常见的嵌入式微控制器一般都是带以太网控制器（MAC 控制器）的，但是一般都不带物理层，原因还是微控制器大多采用数字电路设计，而以太网物理层含有相当的模拟电路部分，不适合集成。不过在图 1-1 中展示的 Z490 PCH 芯片应该不包含 2.5G 以太网控制器，也就是不包括 MAC 层，支持 2500Mbps 的 link 速率的 MAC 控制器和物理层都由英特尔的 I225-

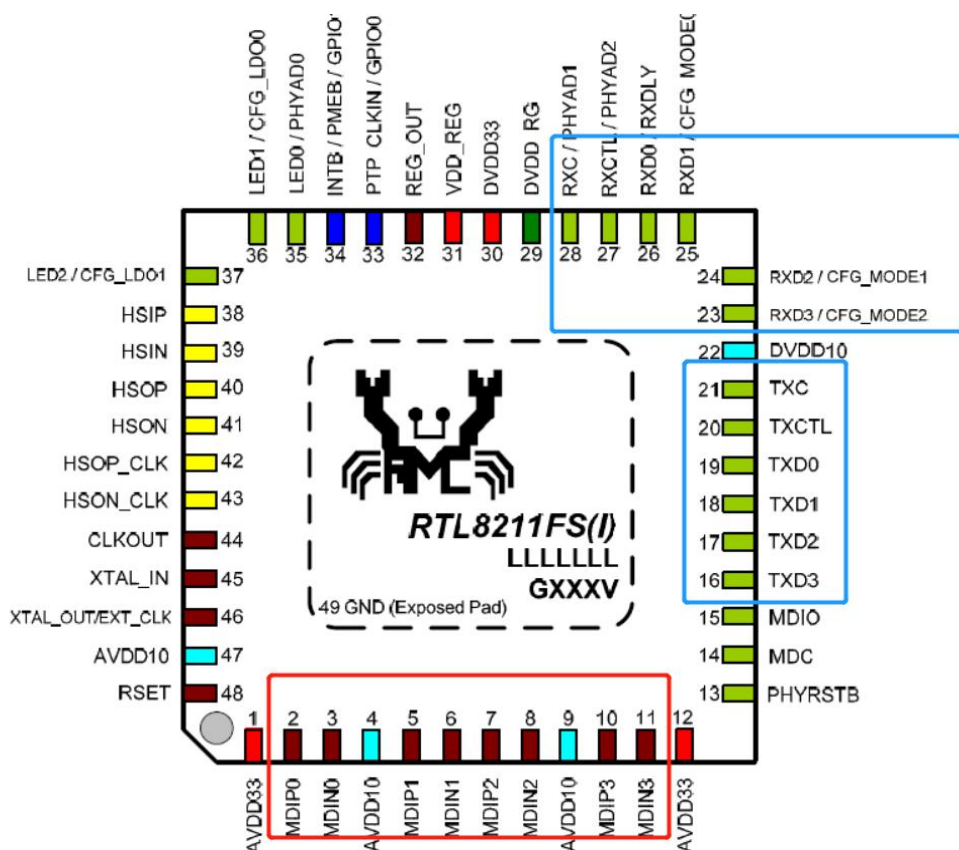
V 网卡芯片实现。这里放两个嵌入式行业实际使用的例子，一是江苏沁恒设计的国产带百兆物理层的 CH563 单片机，它将 MAC 控制器和 PHY 集成在一个芯片上，如图 1-4。

图 1-4 CH563Q 引脚分布图



红圈中就是以太网相关的引脚，其中 RXN,RXP,TXN,TXP 就是 MDI 接口的收发引脚，剩余的是电源和参考电阻引脚。另一个例子是 RTL8211FS，由台湾瑞昱设计的千兆物理层芯片。

图 1-5 RTL8211FS 的引脚分布图



红圈中就是以太网 MDI 接口引脚，和 CH563Q 不同的是，RTL8211FS 的 MDI 接口包含四对差分线，它支持千兆以太网。另外蓝圈中是 MAC 控制器连接物理层的 MII 接口引脚，收发各六个。

物理层芯片主要起编解码的作用，同时它也能把以太网数据帧的有用部分截取下来再转给 MAC 控制器。以太网上的数据是有固定的格式的，和 USB 等高速通讯协议类似，以太网传输数据也是以帧为单位，每帧以同步码开头以同步时钟，之后跟着 SFD，表示数据即将开始，再之后就是有用的数据，而物理层在把以太网信号解码转发给 MAC 控制器的过程中，一般会主动去掉同步码和 SFD。以太网数据帧的具体格式，会在下章的数据链路层讨论。

10BASE-T

在双绞线上，以太网帧以编码好的信号进行传输。下面讨论几种常见的，且使用双绞线做介质的以太网规范。

标准以太网是最先广泛应用的以太网标准，其速度是 10Mbps。10M 以太网（这里讨论使用 10BASE-T 标准的 10M 以太网）使用的是曼切斯特编码，曼切斯特编码的方式是每个周期中间的下降沿代表“0”，上升沿代表“1”。下图是曼切斯特编码的示意图和实际从双绞线上抓取的 10M 以太网波形。

图 1-6 曼斯斯特编码的波形示意图，最下面的编码波形是以太网使用的

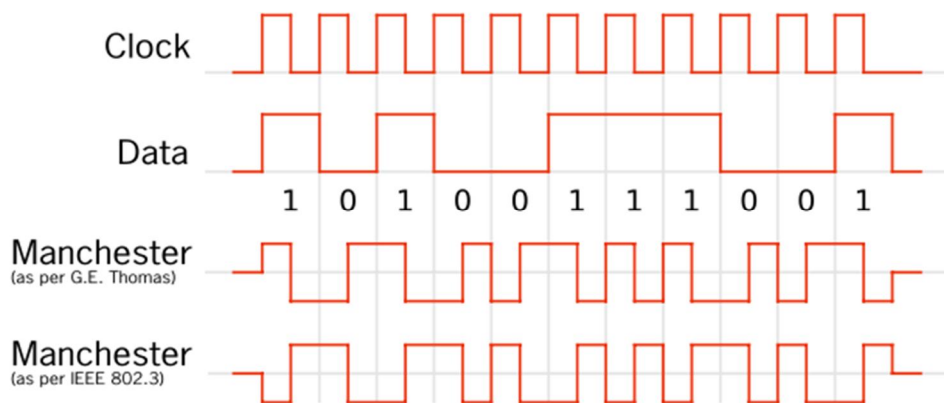
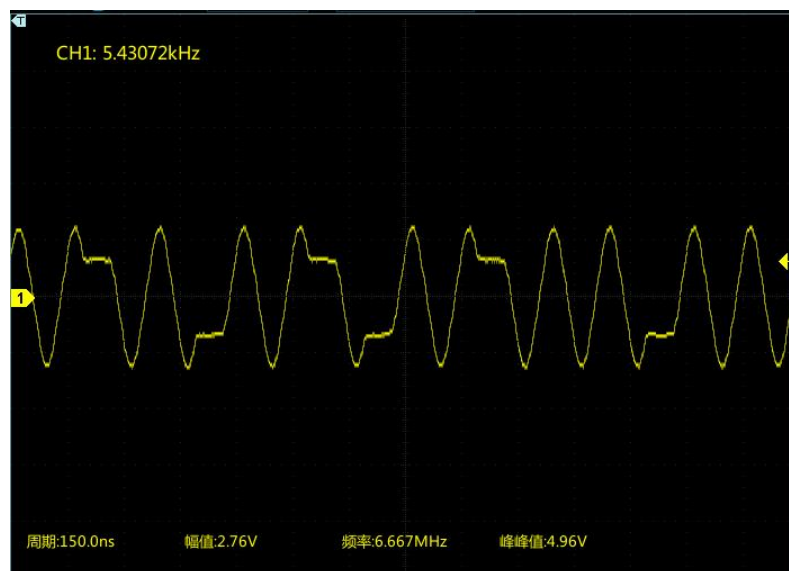


图 1-7 10M 以太网编码波形



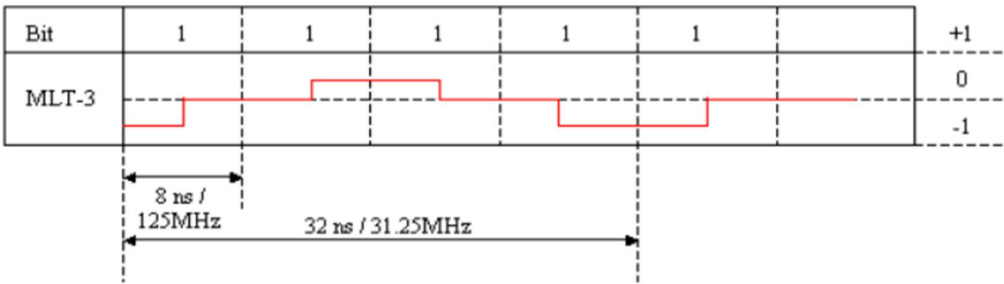
我觉得曼切斯特编码的目的就是为了将时钟和数据调制在一起，并产生足够的跳变，跳变太少不利于交变信号传输。每个曼切斯特编码的信号周期中点必然有一次跳变，这个周期中点的跳变可以被认为是时钟，而以太网帧最开始的同步码可以用来将时钟同步，这样物理层可以很方便地将波形解码成数据。曼切斯特编码也是有缺点的，比如 10M 以太网波形信号只有 10Mbps 的比特率，却要可能会编码成 20MHz 的正弦信号，利用率较低。

100BASE-TX

快速以太网的标准在 1980 年提出，其速度是 100Mbps。100M 以太网（这里讨论用到最广泛的 100BASE-TX 标准）上的波形则不再使用曼切斯特编码。100M 以太网的数据需要先经过 4B5B 编码，也就是每四个比特的数据会先被编码成五个比特，这样 100Mbps 的比特流就变成了 125Mbps。4B5B 编码后的数据，如果使用曼切斯特编码，就会大大超出五类线最大传输 100MHz 带宽的能力。100M 以太网使用的是三电平编码（MLT-3），也就是信号会始终遵循从

零参考点跳到正峰值，再跳到零，然后跳到负峰值，最后跳到零的一个趋势循环。原始数据经过 4B5B 编码，再进行三电平编码，如果是 1 就跳一次，如果是 0 就不跳，这样大大降低了最终介质上波形的频率。当然也不会降低到波形不跳的情况，经过 4B5B 编码的数据最少有 3 个“1”，最多 4 个“1”。不过三电平的编码也导致解码电路略微复杂。下图是 100M 以太网三电平编码方案的示意图、眼图和实际的以太网波形。

图 1-8 100M 以太网三电平编码示意图



“1” = 状态跳变；“0” = 保持状态不变, 使整体信号频率降低.

图 1-9 100M 以太网信道编码眼图

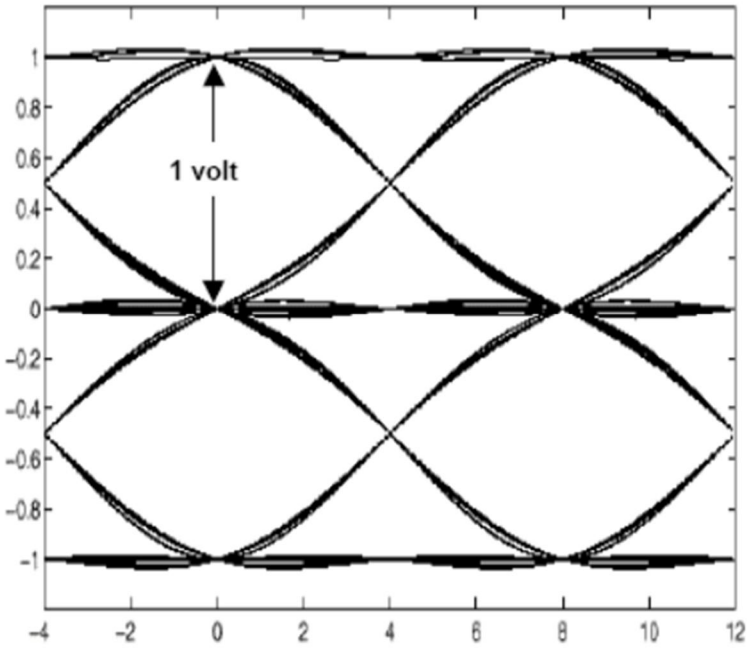
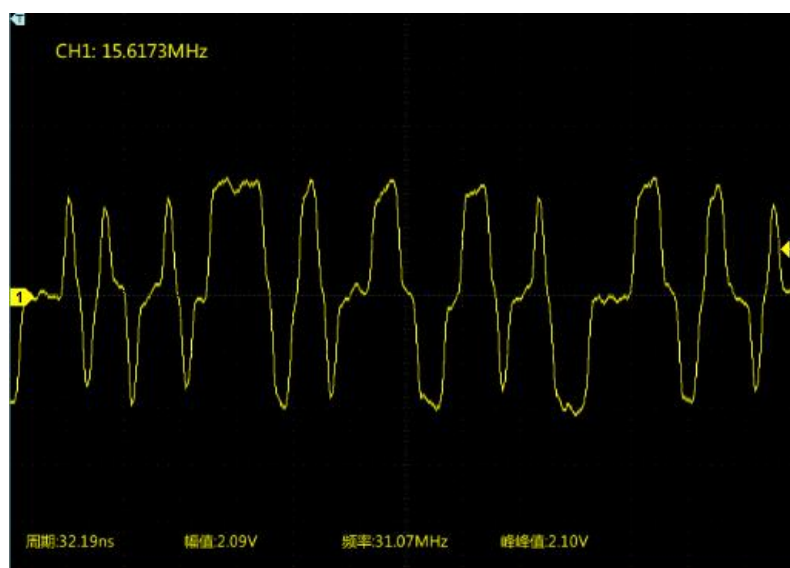


图 1-10 100M 以太网双绞线波形



1000BASE-T

吉比特以太网(IEEE802.3ab),而千兆以太网(这里讨论使用最广泛的 1000BASE-T 标准)的情况就更为复杂。和百兆以太网类似,千兆以太网使用的是五电平(PAM-5)的编码方案,这里为叙述方便起见,将五个电平分别称为: -2, -1, 0, 1, 2 五个值,一般情况下分别对应 -1V/-0.5V/0V/0.5V/1V。千兆以太网不像百兆以太网的电平跳越是朝一个方向连续的,它的跳跃是随机的, -1 电平既可能跳到-0.5, 还可能跳到 1, 它的电平和数据表示关系是,一个电平表示两位数据,即 00, 01, 10 和 11, 还有一个电平用作向前纠错码(FEC)不表示数据。如下图是 1000BASE-T 的编码示意图和信号眼图。

图 1-12 1000BASE-T 编码示意图

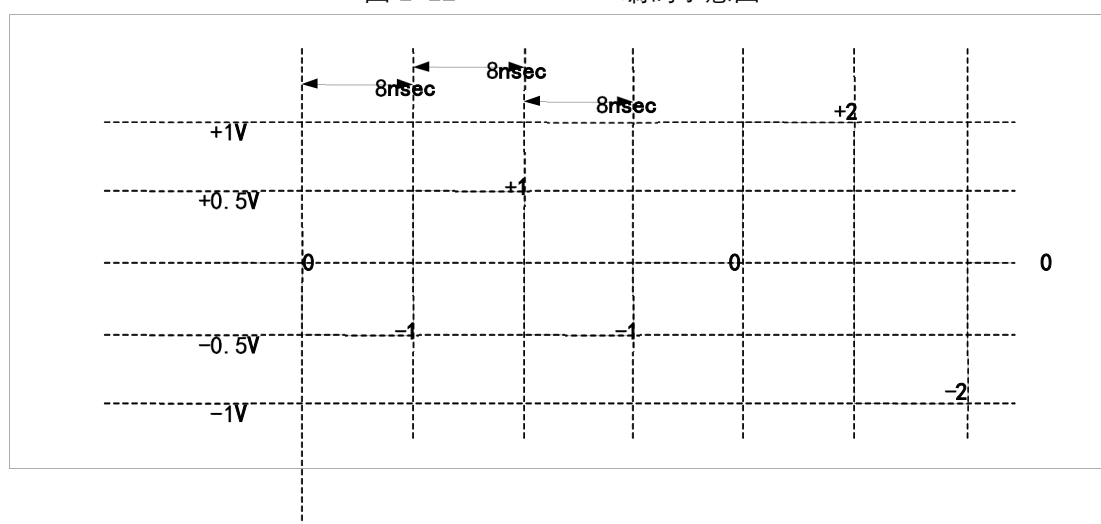
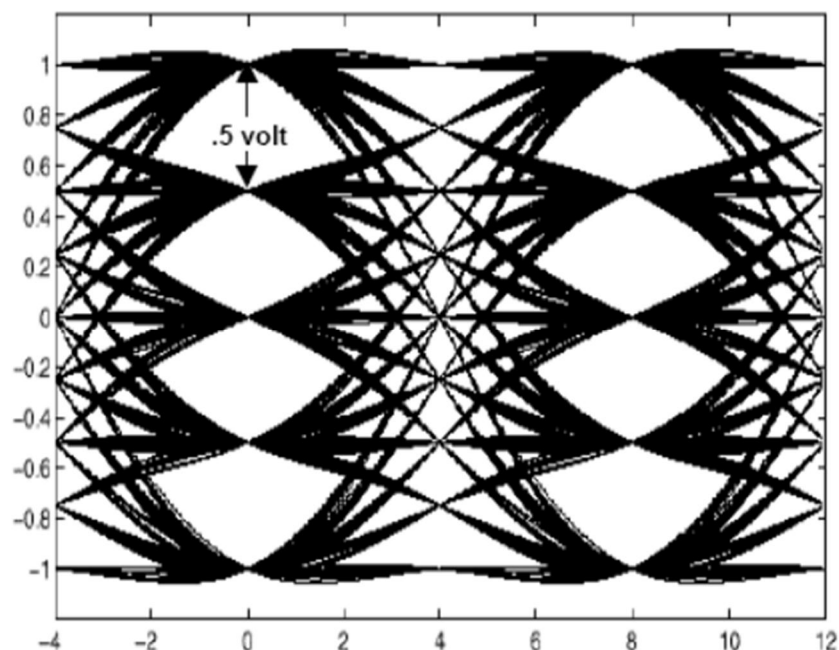


图 1-13 1000BASE-T 信号眼图

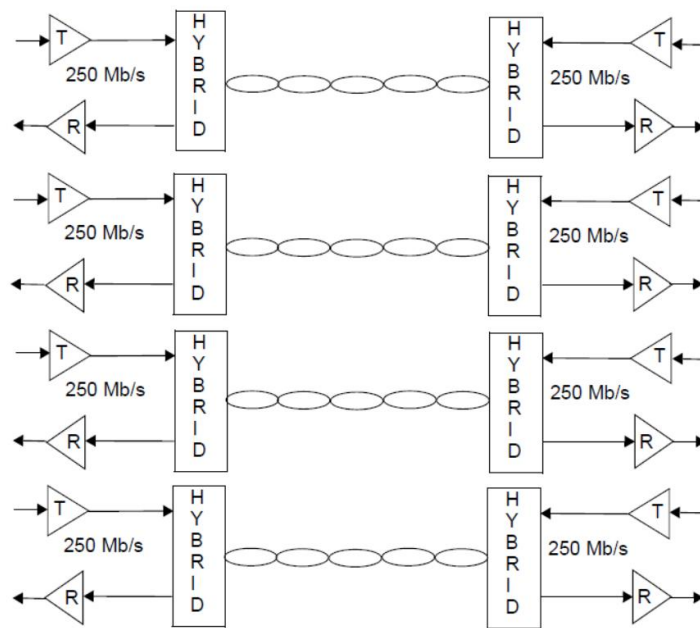


这样看来五电平的编码方案相对于三电平，它的最终频率相对集中一些，不会看起来和三电平一样总是在变。千兆以太网（这里专指 1000BASE-T）使用四对双绞线同时做收发，每对双绞线分担 250Mbps 的速率任务，而由于其采用五电平编码，实际四个电平有效，每个电平代表两位的编码模式，使用 1000BASE-T 的千兆以太网的单个双绞线上的波形频率为 125MHz，勉强能在带宽为 100MHz 的五类线上跑。1000BASE-T 规定信号要能在五类线上传输至少 100 米。

这里提一下另一个实现千兆以太网走双绞线的标准，即 1000BASE-TX。1000BASE-TX 并不是由 IEEE 制定的，而是由美国国家标准协会（ANSI）制定。和百兆以太网的 100BASE-TX 标准类似，1000BASE-TX 固定使用两对双绞线来发送，两外两对双绞线来接收，这样每对差分线将分担 500Mbps 的速率任务，我只查到 1000BASE-TX 采用了 8B10B 编码，但并没有查到其具体采用的波形编码方式，有了解的朋友可以在评论区告诉我一下，先谢过了。可预想的是由于每对双绞线需要实际传输 625Mbps 的信息码率，那么经过编码后的波形带宽怎么也难以控制在五类线的 100MHz 限制下，所以 1000BASE-TX 是需要传输在六类双绞线上的，如此一来就增加了实施成本，所以 1000BASE-TX 的方案没有 1000BASE-T 应用得广泛。

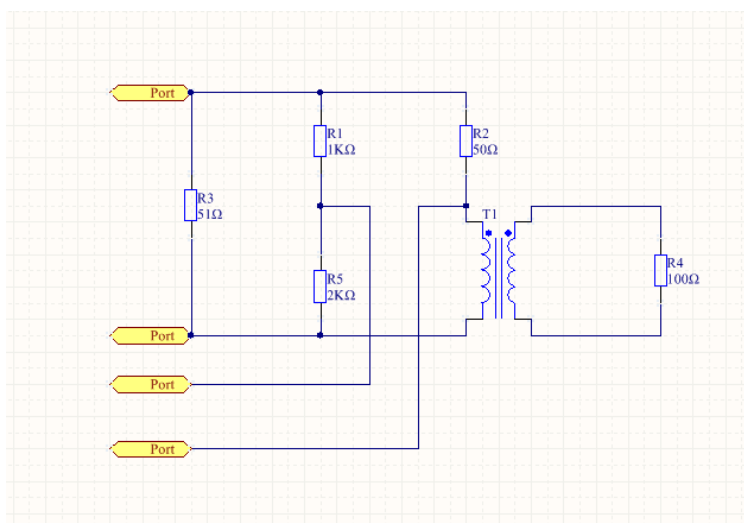
如上面叙述，使用 1000BASE-T 标准的千兆以太网是收发同时使用全部的四对双绞线同时进行收发的，也就是每对双绞线同时进行收发两个方向的操作。这听起来有点扯淡但实际上确实是如此操作的：物理层芯片发出去的信号，不能在自己的接收测引起响应。这需要一个 hybrid 电路，即混波电路。如下图所示。

图 1-14 1000BASE-T 的拓扑结构



这里放一个最简单的混波电路进行混波电路的原理的展示，实际物理层的混波电路会更加复杂，这里就此打住不再继续讨论。各位有兴趣的话我们单独开一章讨论混波电路。

图 1-15 一个简单的混波电路实例



MII 接口

物理层芯片和MAC控制器是通过**介质无关接口**(也有时译作**媒体独立接口**)连接起来的，即 **MII 接口**，如图 2 中提到的 MII 接口就处在物理层和 MAC 控制器之间。之所以命名为介质无关接口，我的理解是 MAC 控制器通过此接口连接物理层芯片，并不关心物理层实际使用的介质类型。在 MAC 允许物理层芯片进行自动协商的情况下，一般 MAC 控制器会通过 MII 接口和物理层沟通，获取实际协商出的速度和连接状态，并设置自身的一些参数。

MII 接口有很多具体的种类，比如，标准 MII 接口，RMII 接口，GMII 接口,RGMII 接口，SGMII 接口等等，名字带“R”表示这是个简化的接口，带“G”表示这个接口支持千兆比特

的通讯速率，带“S”表示这是个串行接口。MII 接口的种类很多，IEEE802.3 定义的 MII 接口只有少部分，很多衍生出来 MII 接口是其他的公司和组织定义的，但市场上也在大量使用。这里详细讨论标准 MII 接口，RGMII 接口，RMII 接口。以下面这个表格说明这三者的关系。

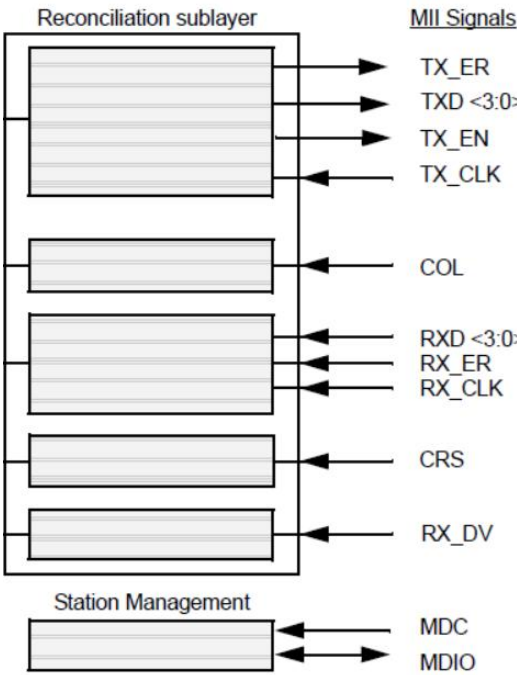
表 1 标准 MII、RMII 和 RGMII 的基本特征

	标准 MII	RMII	RGMII
支持的通讯速率	10M/100M	10M/100M	10M/100M/1000M
发送时钟频率	2.5MHz/25MHz	没有发送时钟	2.5MHz/25MHz/125MHz
发送数据线根数	4	2	4
接受时钟频率	2.5MHz/25MHz	没有接收时钟 CLK_REF 为 50MHz 注[1]	2.5MHz/25MHz/125MHz
接收数据线根数	4	2	4

注[1]: RMII 的时钟一般是由外部产生的，固定为 50MHz，有的 MAC 控制器或者 PHY 芯片会从内部产生 50MHz 通过单独的引脚输出给 RMII 接口使用。

标准 MII 接口在 10/100MH 自适应的 PHY 上十分常见，有发送时钟(TXC)、发送数据使能(TXEN)、发送数据错误(TXER)、发送数据线(TXD[3:0])、接受时钟(RXC)、接受数据有效(RXDV)、接受数据错误(RXER)、接收数据线(RXD[3:0])、载波侦测(CRS)和冲突检测信号(COL)再加上 SMI（stationmanagement interface，站点管理接口）的 MDC 和 MDIO 共 18 根线，如下图所示：

图 1-16 MII 接口（MAC 控制器端视图）



和图中结构呈现不同的是，市面上微控制器的 MAC 控制器的 MDC 时钟一般是由自己发出的。

各线定义如下：

TXC：发送时钟，为 TXD[3:0]、TXER、TXEN 信号提供参考时钟，由 PHY 提供，PHY 在 TXC 信号的上升沿进行数据采样。100M 模式下时钟频率为 25MHz，10M 模式下时钟频率为 2.5MHz；

TXEN：发送数据使能，表示发送的数据有效。TXEN 信号由 MAC 控制器提供，与 TXC 保持同步。

TXER：发送数据错误，TXER 有效代表在此段时间内在 TXD 信号线上传输的数据信号是无效的。TXER 信号由 MAC 控制器提供，与 TXC 保持同步。

TXD[3:0]：发送数据，4 位数据宽度，与 TXC 保持同步。

RXC：接收时钟，为 RXD[3:0]、RXER、RXDV 信号提供参考时钟，也由 PHY 提供，MAC 控制器在 RXC 的上升沿采样数据。100M 模式下时钟频率为 25MHz，10M 模式下时钟频率为 2.5MHz。

RXDV：接收数据有效，由 PHY 驱动，与 RXC 保持同步，表示接收数据有效。

RXER：接收数据错误，由 PHY 驱动，与 RXC 保持同步。RXER 有效且在 RXDV 有效的情况下，表示 RXD 信号线上传输的信号存在错误。

RXD[3:0]：接收数据，4 位数据宽度，由 PHY 提供，与 RXC 保持同步。

CRS：载波侦测，由 PHY 提供，不需要和时钟同步。PHY 工作在半双工模式下时，当接收或发送媒质中任意一个处于非空闲模式下时由 PHY 驱动 CRS 有效，在接收和发送媒质都处于空闲模式时，使 CRS 信号处于无效。

COL：冲突检测，由 PHY 驱动，不和时钟同步，当检测到媒质上有冲突的时候驱动有效，并且保持到冲突结束。

MDC：SMI 接口的时钟。一般由 MAC 发出，MAC 需要保证这个时钟的频率范围是 PHY 所接受的。

MDIO：SMI 接口的数据，是双向接口，一般有 100K 的上拉电阻保证其不被驱动时在高电平状态。MDIO 上的数据有固定的格式。

RMII 接口

RMII 接口即精简的 MII 接口，可以工作在 10M/100M 的连接速率下，收发共用一个 50MHz 的 REF_CLK 时钟，需要由外部产生并输入给 MAC 和 PHY。数据线收发两个方向都改为 2 根，保留数据有效信号，其中接收有效信号和载波检测信号复用。RMII 的引脚一般如下：

TXD[1:0]：发送数据线；

TX_EN：发送数据有效

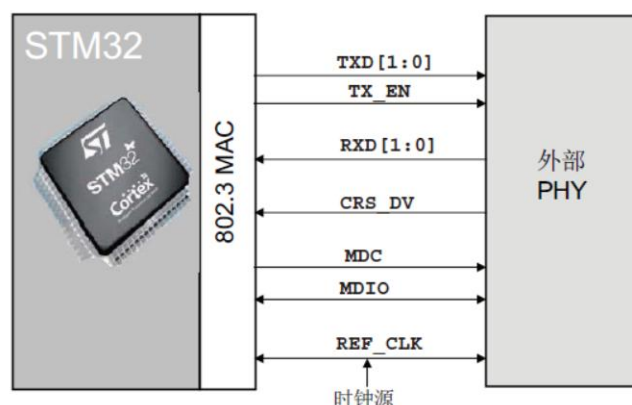
RXD[1:0]:接受数据线;

CRS_DV:载波/接受数据有效引脚;

REF_CLK:50MHz 参考时钟输入, 需要注意的是, 这个引脚在 MAC 控制器这一侧一般是输入引脚, 而有的 PHY 支持从 REF_CLK 引脚输出 50MHz 而有的不支持。为了实现 RMII 接口, 很多 PHY 和 MAC 控制器支持从额外的引脚输出 50Mhz。

SMI 接口的 MDIO 和 MDC。如图所示是 STM32F 系列的 RMII 引脚描述。

图 1-17 STM32F 系列的 RMII 引脚描述

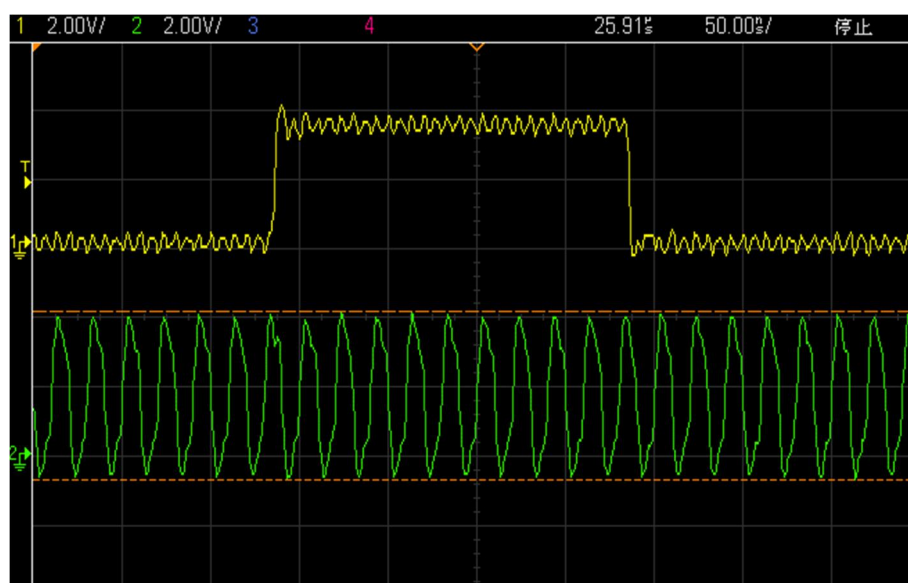


ai15624

一般可以从 STM32 的 MCO 输出引脚输出 50MHz。

RMII 的时钟速率固定为 50MHz, 数据线位宽为 2 位, 所以 RMII 的速率可以达到 100Mbps。但工作 10Mbps 的连接速率时, RMII 每个数据线的每位数据会保留 10 个周期, 从而使比特流的速率降到 10Mbps, 如下图。

图 1-18 RMII 接口工作在 10Mbps 的数据线和时钟线上的波形。



注: 黄色的波形来自一根数据线, 绿色的波形来自时钟线

RGMII 接口

RGMII 接口即精简的千兆位传输速率的 **MII** 接口，是我本章的重点讨论接口，是嵌入式产品常用的千兆以太网 **MII** 接口。同样能达到此速率的接口还有 **GMII**、**SGMII** 等。**RGMII** 也不是 **IEEE802.3** 定义的标准，它是惠普公司定义的。

与标准 **MII** 接口类似，**MII** 接口有收发两组信号线组成，每组有一根时钟线，一根数据有效线，四根数据线组成，时钟线的频率在工作速率分别为 10M/100/1000M 模式时，速率分别为 2.5MHz/25MHz/125MHz，其中当工作在千兆位每秒的速率时，**RGMII** 会在时钟线的上升下降两个跳变沿对数据进行采样，也即是 **DDR** 模式。

RGMII 的各个引脚如下所描述：

TXC，发送时钟，在工作速率分别为 10M/100/1000M 模式时，频率分别为 2.5MHz/25MHz/125MHz，精度要求在 $\pm 50\text{ppm}$ 。

TXCTL，发送数据控制，在时钟上升沿时为 **TXEN**，表示数据有效；时钟下降沿时为 **TXERR**，表示数据错误。

TXD[3:0]，发送数据信号，在时钟上升沿时为 **bit[3:0]**，时钟下降沿时为 **bit[7:4]**。

以上的信号都是由 **MAC** 控制器发送到 **PHY**。

RXC 接收时钟，在工作速率分别为 10M/100/1000M 模式时，频率分别为 2.5MHz/25MHz/125MHz，精度要求在 $\pm 50\text{ppm}$ 。

RXCTL，发送数据控制，在时钟上升沿时为 **RXDV**，表示数据有效；时钟下降沿时为 **RXERR**，表示数据错误。

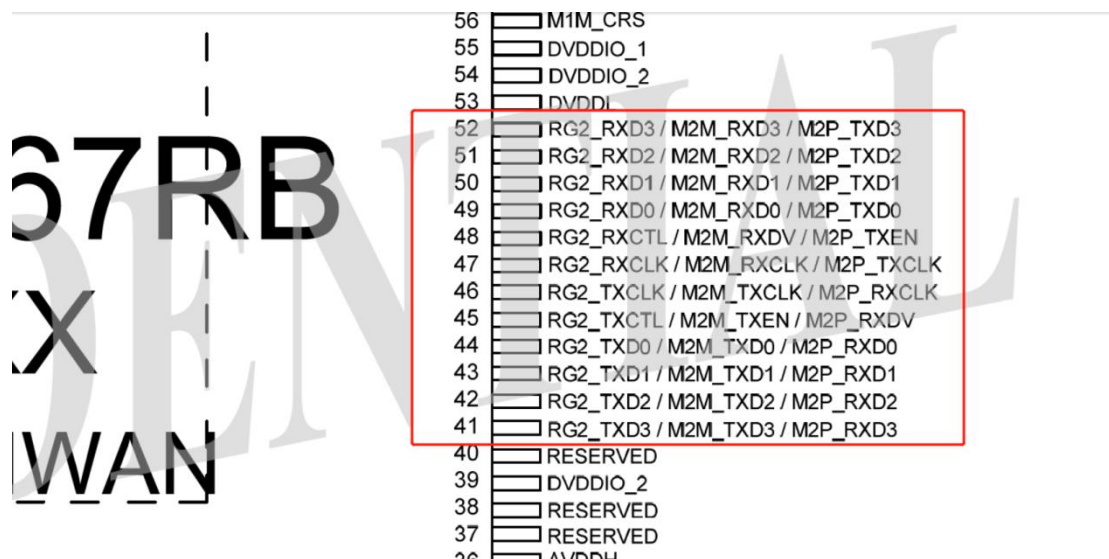
RXD[3:0]，接收数据信号，在时钟上升沿时为 **bit[3:0]**，时钟下降沿时为 **bit[7:4]**。

以上的信号都是由 **PHY** 发送到 **MAC** 控制器发送到。

注意，有的 **MAC** 控制器或者 **PHY** 的 **TXCTL** 和 **RXCTL** 只有指示数据有效的作用，没有指示数据错误的作用。

如图是瑞昱经典的傻瓜交换机芯片 **RTL8367RB** 的 **RGMII** 接口。

图 1-19 **RTL8367RB** 的 **RGMII** 接口（**RG2_**开头的引脚）



由于采用双边沿采样的模式，因此数据线之间的等长情况更需要注意，一般应该把数据线的长度差控制在 9.8mm，时钟线上时钟到达对侧的时间要比数据线延迟 1.5ns 到 2ns 之间，走线的特性阻抗建议保持 50Ω。需要注意的是，有的 MAC 和有的 PHY 在芯片内部做了时钟相对数据延迟 2ns 的操作，那么外部的走线只需要保证等长即可。

本章总结

本章具体讨论了如何在嵌入式设备的层面实现以太网物理层，以及物理层编解码芯片，MDI 和 MII 接口，10BASE-T/100BASE-TX/1000BASE-T 介质层面的标准。