

以太网 PHY 寄存器分析

以太网 PHY 寄存器分析.....	1
1、以太网 PHY 标准寄存器分析.....	2
1.1 Control Register	2
1.2 Status register.....	5
1.3 PHY Identifier Register.....	8
1.4 Auto-Negotiation Advertisement Register.....	8
1.5 Auto-Negotiation Link Partner Base Page Ability Register.....	9
1.6 Auto-Negotiation Expansion Register	10
1.7 AN next page Register/AN Link Partner Received Next Page	10
1.8 MASTER-SLAVE Control Register.....	10
1.9 MASTER-SLAVE Status Register	12
1.10 Extended Status Register.....	13
2、PHY 扩展寄存器分析.....	13
2.1 工作模式控制器.....	14
2.2 端口驱动模式.....	15
2.3 预加重配置.....	15
2.4 自动协商降格.....	16
2.5 Auto-Crossover 配置	17
2.6 MDI 信号边沿速率调整.....	18
2.7 错误指示寄存器.....	18

1、以太网 PHY 标准寄存器分析

PHY 是 IEEE802.3 中定义的一个标准模块，STA (station management entity, 管理实体，一般为 MAC 或 CPU) 通过 SMI (Serial Manage Interface) 对 PHY 的行为、状态进行管理和控制，而具体管理和控制动作是通过读写 PHY 内部的寄存器实现的。PHY 寄存器的地址空间为 5 位，从 0 到 31 最多可以定义 32 个寄存器 (随着芯片功能不断增加，很多 PHY 芯片采用分页技术来扩展地址空间以定义更多的寄存器，在此不作讨论)，IEEE802.3 定义了地址为 0-15 这 16 个寄存器的功能，地址 16-31 的寄存器留给芯片制造商自由定义，如表 1 所示。以下结合实际应用，对 IEEE802.3 定义的寄存器各项功能进行分析。

表 1 PHY 管理寄存器集

Register address	Register name	Basic/Extended	
		MII	GMII
0	Control	B	B
1	Status	B	B
2,3	PHY Identifier	E	E
4	Auto-Negotiation Advertisement	E	E
5	Auto-Negotiation Link Partner Base Page Ability	E	E
6	Auto-Negotiation Expansion	E	E
7	Auto-Negotiation Next Page Transmit	E	E
8	Auto-Negotiation Link Partner Received Next Page	E	E
9	MASTER-SLAVE Control Register	E	E
10	MASTER-SLAVE Status Register	E	E
11 through 14	Reserved	E	E
15	Extended Status	Reserved	B
16 through 31	Vendor Specific	E	E

1.1 Control Register

寄存器 0 是 PHY 控制寄存器，通过 Control Register 可以对 PHY 的主要工作状态进行设置。Control Register 的每一位完成的功能见表 2。

表 2 Control Register

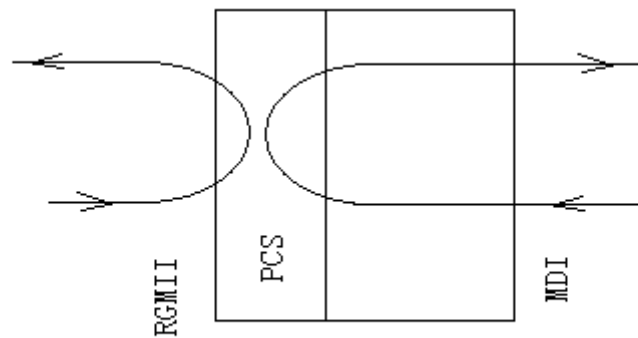
Bit(s)	Name	Description	R/W/a
0.15	Reset	1 = PHY reset 0 = normal operation	R/W SC
0.14	Loopback	1 = enable loopback mode 0 = disable loopback mode	R/W
0.13	Speed Selection (LSB)	0.6 0.13 1 1 = Reserved 1 0 = 1000 Mb/s 0 1 = 100 Mb/s 0 0 = 10 Mb/s	R/W
0.12	Auto-Negotiation Enable	1 = Enable Auto-Negotiation Process 0 = Disable Auto-Negotiation Process	R/W
0.11	Power Down	1 = power down 0 = normal operation	R/W
0.10	Isolate	1 = electrically Isolate PHY from MII or GMII 0 = normal operation	R/W
0.9	Restart Auto-Negotiation	1 = Restart Auto-Negotiation Process 0 = normal operation	R/W SC
0.8	Duplex Mode	1 = Full Duplex 0 = Half Duplex	R/W
0.7	Collision Test	1 = enable COL signal test 0 = disable COL signal test	R/W
0.6	Speed Selection (MSB)	0.6 0.13 1 1 = Reserved 1 0 = 1000 Mb/s 0 1 = 100 Mb/s 0 0 = 10 Mb/s	R/W
0.5:0	Reserved	Write as 0, ignore on Read	R/W

Reset: Bit15 控制的是 PHY 复位功能，在该位置写入 1 实现对 PHY 的复位操作。复位后该端口 PHY 的其他控制、状态寄存器将恢复到默认值，每次 PHY 复位应该在 0.5s 的时间内完成，复位过程中 Bit15 保持为 1，复位完成之后该位应该自动清零。一般要改变端口的工作模式（如速率、双工、流控或协商信息等）时，在设置完相应位置的寄存器之后，需要通过 Reset 位复位 PHY 来使配置生效。

Loopback: Loopback 是一个调试以及故障诊断中常用的功能，Bit14 置 1 之后，PHY 和外部 MDI 的连接在逻辑上将被断开，从 MAC 经过 MII/GMII（也可能是其他的 MAC/PHY 接口）发送过来的数据将不会被发送到 MDI 上，而是在 PHY 内部（一般在 PCS）回环到本端口的 MII/GMII 接收通道上，通过 Loopback 功能可以检查 MII/GMII 以及 PHY 接口部分是否工作正常，对于端口不通的情况可用于故障定位。需要注意的是，很多时候 PHY 设置 Loopback 后端口可能就 Link down 了，MAC 无法向该端口发帧，这时就需要通过设置端口 Force Link up 才能使用 Loopback 功能。

案例：在 S3760-12SFP/GT 开发过程中，我们曾经出现过一个故障，其中有一片 PHY（88E1145）对应的端口发送的帧出现 CRC 错误，当时这个问题的排查过程经历和很长的时间，最后得出的结论是 RGMII 的接口电平配置电阻焊接混料导致故障。我们姑且不去考虑这个案例实际的解决过程，在这里讨论一下如何通过 Loopback 功能对该问题进行定位。

首先介绍一下 S3760 交换部分的架构，MAC 芯片为 98EX126，通过 RGMII 接口连接到 PHY 芯片 88E1145，MAC 通过 PCI 管理总线连接到 CPU。在这个案例中，查看 88E1145 的资料，其 Loopback 操作在 PCS 子层完成，两个方向的 Loopback，如下图所示。第一种模式，从 MAC 经过 RGMII 发送的帧到达 PCS 后被 Loopback 到 RGMII 的接收通道再送回给 MAC（这种模式就是上面所描述的寄存器 0 Loopback 位控制的 Loopback 模式），另一种模式，从 MDI 接收上来的帧到达 PCS 后被 Loopback 到 MDI 的发送通道，这种 Loopback 模式在 IEEE802.3 中并没有要求，但是目前常见的 PHY 都支持该功能。分别做这两种 Loopback 操作，可以发现第一种 Loopback 操作之后可以在 MAC 上检测到 CRC 错误，而第二种 Loopback 模式，用 SMB 从端口砸帧再 Loopback 回来没有检测到 CRC 错误，这样我们就可以判断故障应该在 PCS 以上的部分，并且，两种 Loopback 模式下 PHY 的 PCS 都有再工作，基本上也可以排除 PCS 的故障。因此可以进一步定位到故障在 PHY 的 RGMII 或者 MAC 上。我们就可以去检查这些部分的相关设计来解决问题了。



88E1145 Loopback示意图

要进一步更精确的定位问题，我们还可以去查询 MAC 芯片是否有类似的端口 Loopback 功能，如果有则在 MAC 内部也做一下 Loopback 观察是否有 CRC；如果没有，可以将 MAC 和 PHY 的 RGMII 接口断开，将 MAC 的 RGMII 发送和接收通道自己连接起来，将 PHY 的 RGMII 发送和接收通道自己连接起来，分别做砸帧测试观察有没有 CRC，这样就可以进一步的缩小范围。不过这个 S3760 的案例有其特殊性，98EX126 没有端口的 Loopback 功能，而 MAC 的 RGMII 发送信号直接连接到 PHY，中间没有电阻，而且两者都是 BGA 封装，这两个实验都没办法进行。因此故障排查中需要检查的范围就比较广一点了。但是从中我们可以看出，Loopback 操作在故障定位中可以起到将各个功能模块隔离定位的作用，虽然这些模块在物理上是集成在一个芯片中的。这种分割隔离的思想在故障定位中是非常重要的。

Speed Selection: Bit13 和 Bit6 两位联合实现对端口的速率控制功能，具体的对应关系详见表 2。需要注意的是 Speed Selection 只有在自动协商关闭的情况下才起作用，如果自动协商设置为 Enable 状态，则该设置不起作用；并且，对 Speed Selection 的修改设置，往往需要复位端口才能配置生效。因此在设置该位置的时候需要检查自动协商的设置并通过 Bit15 复位端口。

Auto-Negotiation Enable: 自动协商（AN）开关。设置为 1 表示打开 AN 功能，端口的工作模式通过和连接对端进行 AN 来确定。如果设置为 0 则 AN 功能

关系，端口的工作模式通过 Control Register 相应位置的配置决定。必须注意的是，对于 1000BASE-T 接口,自动协商必须打开。

Power Down: 端口工作开关。设置为 1 将使端口进入 Power Down 模式，正常情况下 PHY 在 Power Down 模式其 MII 和 MDI 均不会对外发送数据。Power Down 模式一般在软件 shut down 端口的时候使用，需要注意的是端口从 Power Down 模式恢复，需要复位端口以保证端口可靠的连接。

Isolate: 隔离状态开关。改位置 1 将导致 PHY 和 MII 接口之间处于电气隔离状态，除了 MDC/MDIO 接口的信号外，其他 MII 引脚处于高阻态。IEEE802.3 没有对 Isolate 时 MDI 接口的状态进行规范，此时 MDI 端可能还在正常运行。Isolate 在实际应用中并没有用到。并且，值得注意的是，由于目前很多百兆的 PHY 芯片其 MAC 接口主流的都是 SMII/S3MII，8 个端口的接口是相互关联的，一个端口设置 Isolate 可能会影响其他端口的正常使用，因此在使用中注意不要随意更改 bit10 的状态。

Restart Auto-Negotiation : 重新启动自动协商开关。Bit9 置 1 将重新启动端口的自动协商进程，当然前提是 Auto-Negotiation Enable 是使能的。一般在修改端口的自动协商能力信息之后通过 Bit9 置 1 重新启动自动协商来使端口按照新的配置建立 link。

Duplex Mode: 双工模式设置。Bit8 置 1 端口设置为全双工，置 0 则端设置为半双工，和 Speed Selection 的设置一样，Duplex Mode 的设置只有在自动协商关闭的情况下才起作用，如果自动协商设置为 Enable 状态，则该设置不起作用，端口的双工模式根据 AN 结果来定。对 Duplex Mode 的修改配置也需要复位端口才能生效。

Collision Test: 冲突信号 (COL) 测试开关。在需要对 COL 信号进行测试时，可以通过 Bit7 置 1，这时 PHY 将输出一个 COL 脉冲以供测试。实际测试操作中也可以将端口配置为半双工状态，通过发帧冲突来测试 COL 信号，因此该配置实用价值不大。

1.2 Status register

寄存器 1 是 PHY 状态寄存器，主要包含 PHY 的状态信息，大多数 bit 的值都是由芯片厂家确定的，每一个 bit 的功能在表 3 种已有详细说明。其中指示 PHY 所具有的工作模式能力的寄存器不再多讲，值得注意的有以下几位。

表 3 Status register

Bit(s)	Name	Description	R/Wa
1.15	100BASE-T4	1 = PHY able to perform 100BASE-T4 0 = PHY not able to perform 100BASE-T4	RO
1.14	100BASE-X Full Duplex	1 = PHY able to perform full duplex 100BASE-X 0 = PHY not able to perform full duplex	RO
1.13	100BASE-X Half Duplex	1 = PHY able to perform half duplex 100BASE-X 0 = PHY not able to perform half duplex	RO

100BASE-X

1.12	10 Mb/s Full Duplex	1 = PHY able to operate at 10 Mb/s in full duplex mode 0 = PHY not able to operate at 10 Mb/s in full duplex mode	RO
1.11	10 Mb/s Half Duplex	1 = PHY able to operate at 10 Mb/s in half duplex mode 0 = PHY not able to operate at 10 Mb/s in half duplex mode	RO
1.1	100BASE-T2 Full Duplex	1 = PHY able to perform full duplex 0 = PHY not able to perform full duplex	RO
1.9	100BASE-T2 Half Duplex	1 = PHY able to perform half duplex 0 = PHY not able to perform half duplex	RO
1.8	Extended Status	1 = Extended status information in Register 15 0 = No extended status information in Register 15	RO
1.7	Reserved	ignore when read	RO
1.6	MF Preamble Suppression	1 = PHY will accept management frames with preamble suppressed. 0 = PHY will not accept management frames with preamble suppressed.	RO
1.5	Auto-Negotiation Complete	1 = Auto-Negotiation process completed 0 = Auto-Negotiation process not completed	RO
1.4	Remote Fault	1 = remote fault condition detected 0 = no remote fault condition detected	RO/ LH
1.3	Auto-Negotiation Ability	1 = PHY is able to perform Auto-Negotiation 0 = PHY is not able to perform Auto-Negotiation	RO
1.2	Link Status	1 = link is up 0 = link is down	RO/ LL
1.1	Jabber Detect	1 = jabber condition detected 0 = no jabber condition detected	RO/ LH
1	Extended Capability	1 = extended register capabilities 0 = basic register set capabilities only	RO

Auto-Negotiation Complete: AN 完成状态指示位。Bit5 指示的是端口 AN 进程是否完成的状态位。在 AN Enable 的情况下，Bit5=1 表示自动协商进程已经成功结束，此时 PHY 的其他和 Link 状态相关的寄存器才是正确可靠的。如果 AN

进程没有完成，则这些状态信息可能是错误的。在调试以及异常故障处理时，可以通过该位寄存器的状态判断 AN 是否成功，从而进一步的检查 AN 相关的设置是否正确，或者芯片的 AN 功能是否正常等。

Remote Fault: 远端错误指示位。Bit4=1 代表连接对端 (Link Partner) 出错，至于出错的具体类型以及错误检测机制在规范中并没有定义，由 PHY 的制造商自由发挥，一般的厂商都会在其他寄存器 (Register16-31 由厂商自行定义) 指示比较详细的错误类型。在与端口相关的故障查证中，Remote Fault 是一个重要的指示信息，通过互联双方的 Remote Fault 信息 (可能要加上其他的具体错误指示)，可以帮助定位故障原因。

Link Status: Link 状态指示位。Bit2=1 代表端口 Link up, 0 则代表端口 Link down。实际应用中一般都是通过 Bit2 来判断端口的状态。而且，一般的 MAC 芯片也是通过轮询 PHY 的这个寄存器值来判断端口的 Link 状态的 (这个过程可能有不同的名称，比如 BCM 叫做 Link Scan, 而 Marvell 叫做 PHY Polling。) 如前所述，在 AN Enable 的情况下，Link Status 的信息只有在 Auto-Negotiation Complete 指示已经完成的情况下才是正确可靠的，否则有可能出错。

案例: 曾经有发现过一个故障，我司 S3760 的 SFP 端口和 Cisco 设备互联，发现端口 Link 指示灯已经点亮，但是软件显示的端口状态却是 Link down，并且端口也不能转发帧。读取 S3760 的 PHY 寄存器，发现 Link Status=1，而读 MAC 的状态寄存器，发现其 Link 状态位为 0，软件就是据此判断端口为 Link down 的。可以看出，故障的直接原因是 MAC 和 PHY 的 Link 状态不一致。但是为什么 MAC 和 PHY 状态不一致呢？读取 Auto-Negotiation Complete 状态指示寄存器，发现 Auto-Negotiation Complete=0，显然自动协商没有完成。检查互联双方的端口配置，我司 S3760 的配置为 AN Enable，而 Cisco 设备 AN 是 Disable 的。这样的配置显然自动协商不可能完成，将我司 S3760 的端口也配置为 AN Disable 的强制状态，端口即可以正常 Link Up 和转发帧了。同时据此信息向芯片制造商方面咨询，对方的答复是，PHY Polling 查询 PHY 状态时，如果端口为 AN Enable，则一定要等待 Auto-Negotiation Complete=1，才认为 PHY 的 Link status 有效。这就可以解释为什么 MAC 和 PHY 的 Link 状态不一样了。

但是，为什么 PHY 的在 AN 尚未完成的时候 Link status 就已经置 1 了呢？原来 3760 的 PHY 有一个配置，1000BASE-X AN Bypass 功能，PHY 如果在 AN 过程中没有收到对方的 AN 信息，则可以跳过 AN 进程，通过检测 Serdes 接口上的信号来建立 Link。这本来是一个很好的特色功能，可是由于 PHY 通过 1000BASE-X AN Bypass 功能来建立 Link 之后却没有将 Auto-Negotiation Complete 位置 1，和 MAC 的 PHY Polling 进程矛盾了，导致 MAC 和 PHY 的 Link 状态不一致。(大家可以实际尝试一下，电口的自动协商同时还定义了一个 Parallel Detect 功能，可以让一个 AN Enable 的端口和一个 AN Disable 的端口建立 Link，但是 PHY 通过 Parallel Detect 建立 Link 其 Auto-Negotiation Complete 位是置 1 的。) 至此，解决的办法就是关闭 PHY 的 1000BASE-X AN Bypass 功能，故障就解决了。

Jabber Detect: Jabber 检测指示位。IEEE802.3 对 Jabber 的解释是“A condition wherein a station transmits for a period of time longer than the maximum permissible packet length, usually due to a fault condition”。这一位指示的是 Link Partner 发送的时间超过了规定的最大长度。值得注意的是，Jabber Detect 只有在 10BASE-T 模式下才有意义，100 和 1000M 模式是没有定义 Jabber 这一功能的。

1.3 PHY Identifier Register

寄存器 2 和 3 存放 PHY 芯片的型号代码，由芯片制造商自行定义，实际应用中软件通过读取这两个寄存器的内容可以识别 PHY 的型号和版本，这些内容都是只读寄存器，对 PHY 的功能没有影响，也不反映 PHY 的工作状态，实用价值不大。

1.4 Auto-Negotiation Advertisement Register

寄存器 4 是自动协商的能力通告寄存器，在 AN Enable 的前提下（见寄存器 0），端口根据该寄存器的相关配置将自动协商信息通过 FLP 在 MDI 上进行通告。当 AN 配置为 Disable 状态的时候，寄存器 4 的配置将不起作用，端口的工作模式由控制寄存器中的配置决定。寄存器 4 的详细定义对电口和光口 PHY 上有不同的定义，其中电口 PHY 的具体说明如表 4A。每个 bit 的功能已有详细描述，无需赘述。

表 4A Auto-Negotiation Advertisement Register (Copper)

Bit(s)	Name	Description	R/W
4.15	Next Page	0=Next Page ability is not supported/No NP to exchange 1=Next Page to exchange	R/W
4.14	Reserved	Write as zero, ignore on read	RO
4.13	Remote Fault	0=don't transmit Remote Fault Information 1=transmit Remote Fault Information	R/W
4.12:5	Technology Ability Field	Technologies supported by local PHY to Advertise	R/W
4.4:0	Selector Field	the type of message being sent by Auto-Negotiation	R/W

Bit12: 5 对应自动协商广播能力域 (Technology Ability Field)，每一位分别对应为 A[7: 0]，每一位配置一种工作模式的能力。在实际应用中，如果 PHY 要支持该种工作模式则对应位置 1，若不支持则对应位置 0。注意到在这 8 位能力指示域中，并没有 1000BASE-T 能力的对应配置位，1000BASE-T 的相关配置在寄存器 9，MASTER-SLAVE Control Register 来完成。

Bit4: 0 配置自动协商的类型，规范正在发送的自动协商信息遵从何种规范，我们所接触的以太网 PHY 遵从 IEEE802.3 规范，Selector Field=0001，该区域不可随意更改（很多 PHY 将此区域设计为只读寄存器，以免被修改）。

Technology Ability Field

Bit	Technology	Minimum cabling requirement
A0	10BASE-T	Two-pair category 3
A1	10BASE-T full duplex	Two-pair category 3
A2	100BASE-TX	Two-pair category 5

A3	100BASE-TX full duplex	Two-pair category 5
A4	100BASE-T4	Four-pair category 3
A5	PAUSE operation for full duplex links	Not applicable
A6	Asymmetric PAUSE operation for full duplex Links	Not applicable
A7	Reserved for future technology	

思考： 在一个交换机端口上配置 Speed 100; Duplex Full; Floecontrol Auto, 请问这时候 Register 0 和 Regiter 4 的值应该分别是多少？

光口 PHY 在这里特指千兆光口（1000BASE-X）的 PHY，其自动协商通告寄存器如表 4B 所述。需要注意的是，1000BASE-X 的 AN 除了双工和流控信息之外，并不能协商速率信息，也就是说端口只能工作在 1000M 模式下。并且，端口的媒介类型（LX/SX）也不能通过自动协商来解决，因此在应用上必须人工保证互联双方的速率、媒介类型的一致性，否则结果将是连接失败，AN 对此无能为力。

表 4B Auto-Negotiation Advertisement Register（1000BASE-X）

Bit(s)	Name	Description	R/W
		0=Next Page ability is not supported/No NP to exchange	
4.15	Next Page	1=Next Page to exchange	R/W
4.14	Reserved	Write as zero, ignore on read	RO
4.13:12	Remote Fault	0=don't transmit Remote Fault Information 1=transmit Remote Fault Information	R/W
4.11:9	Reserved	Write as zero, ignore on read	RO
4.8:7	Pause	0= don't Advertise Pause capability 1= Advertise Pause capability	R/W
4.6	Half Duplex	0= don't Advertise 1000BASE-X HD capability 1= Advertise 1000BASE-X HD capability	R/W
4.5	Full Duplex	0= don't Advertise 1000BASE-X FD capability 1= Advertise 1000BASE-X FD capability	R/W
4.4:0	Reserved	Write as zero, ignore on read	RO

1.5 Auto-Negotiation Link Partner Base Page Ability Register

寄存器 5 保存的是本端 PHY 接收到的对端 PHY 所通告的端口能力，寄存器 5 的结构和寄存器 4 基本一致。应用上，寄存器 5 可以用于检测 Link partner 的自动协商配置，在端口 Link 故障的定位排查中可以发挥重要作用。特别是当 Link Partner 不是我司设备的时候，其内部寄存器信息我们是无法获取的，这时候就只能通过寄存器 5 来获取对方的自动协商信息了。不单单是 AN 信息，端口的状态信息中所有关于 Link Partner 状态的指示信息在我们进行故障处理的时候都是很珍贵的第一手资料，通过分析这些信息对我们进行故障定位将有很大的帮助。

1.6 Auto-Negotiation Expansion Register

寄存器 6 保存了 PHY 自动协商过程的异常信息，每一位的作用在表 5 中一目了然。从这个寄存其中我们可以获取到 Link Partner 子否支持自动协商以及自动协商下一页有没有收到的信息。其中 Parallel Detection Fault 表示，端口在并行检测进程中出现了错误，这包含了两层意义：首先 PHY 已经启动并行检测，则 Linkpartner 不支持 AN，再则并行检测不能成功的探测到 Linkpartner 的连接速率信息。

另外，光口（1000BASE-X）PHY 的这个寄存器只定义了 bit1 和 bit2 两位，含义和电口相同，见下表。

表 5 Auto-Negotiation Expansion Register

Bit(s)	Name	Description	R/W
6.15:5	Reserved	Write as zero, ignore on read	RO
6.4	Parallel Detection Fault	1 = fault detected via the Parallel Detection function. 0 = No fault detected via the Parallel Detection function.	RO/ LH
6.3	Link Partner Next Page Able	1 = Link Partner is Next Page able 0 = Link Partner is not Next Page able	RO
6.2	Next Page Able	1 = Local Device is Next Page able 0 = Local Device is not Next Page able	RO
6.1	Page Received	1 = A New Page has been received 0 = A New Page has not been received	RO/ LH
6.0	Link Partner AN Able	1 = Link Partner is Auto-Negotiation able 0 = Link Partner is not Auto-Negotiation able	RO

1.7 AN next page Register/AN Link Partner Received Next Page

寄存器 7 和 8 分别保存了 Local PHY 和 Linkpartner 的自动协商下一页信息，AN 的下一页功能通常在 1000M 模式的自动协商下使用，详细地寄存器信息要结合 PHY 芯片的资料进行分析，本文不作详细讨论。

1.8 MASTER-SLAVE Control Register

寄存器 9 保存的是 1000BASE-T 模式的配置信息，控制 PHY 的 AN 信息中与 1000BASE-T 相关的协商信息，以及 PHY 在 1000BASE-T 模式下的工作模式。详细信息见表 6。

表 6 MASTER-SLAVE Control Register

Bit	Name	Description	Type
-----	------	-------------	------

9.15:13	Test mode bits	Transmitter test mode operations	R/W
	MASTER-SLA	1=Enable MASTER-SLAVE Manual configuration value	
9.12	VE Manual Config Enable	0=Disable MASTER-SLAVE Manual configuration value Default bit value is 0.	R/W
	MASTER-SLA	1=Configure PHY as MASTER during MASTER-SLAVE negotiation	
9.11	VE Config Value	0=Configure PHY as SLAVE during MASTER-SLAVE negotiation	R/W
	Port type	Indicate the preference to operate as MASTER (multiport device) or as SLAVE (single-port device) if the bit 9.12, is not set.	R/W
	1000BASE-T Full Duplex	1 = Advertise PHY is 1000BASE-T full duplex capable. 0 = Advertise PHY is not 1000BASE-T full duplex capable.	R/W
	1000BASE-T Half Duplex	1 = Advertise PHY is 1000BASE-T half duplex capable. 0 = Advertise PHY is not 1000BASE-T half duplex capable.	R/W
9.7:0	Reserved	Write as 0, ignore on read.	R/W

Test mode bits: 测试模式控制器。默认配置为 000，代表 PHY 处于正常工作模式。写入其他数值则 PHY 进入 Test 模式，在不同的 Test 模式下 PHY 在 MDI 上发送不同类型的信号，以供我们对 PHY 的发送信号进行评估测试。关于测试模式的详细描述见 IEEE802.3 Clause 40.6.1.1.2。

MASTER-SLAVE Manual Config Enable: MASTER-SLAVE 强制配置使能位。1000BASE-T 运行模式下，互连双方的工作模式必须是一端 Master 另一端 Slave，一般情况下在 AN 进程中互联双方会自动协商出一端 Master 另一端 Slave。强制的配置则在 AN 的时候不对 MASTER-SLAVE 信息进行协商，PHY 根据强制的 MASTER-SLAVE 配置进行工作。这样带来的问题是如果互联双方的配置一样（都是 MASTER 或者 SLAVE）则不能 Link up，或者 Link up 之后也不能正常进行数据收发操作。因此实际应用中最好不要使用强制配置。关于 MASTER 和 SLAVE 模式的差异，详见 IEEE802.3 Clause 40 的相关描述。

MASTER-SLAVE Config Value: MASTER-SLAVE 强制配置信息位，在 bit11=1 的情况下，bit12=1 则 PHY 只能工作于 Master 模式，bit12=0 则 PHY 只能工作于 SLAVE 模式。

Port type: 端口模式控制位。Bit10 控制端口在 AN 进程中的 MASTER-SLAVE 优先级，1 代表 MASTER 优先，1 代表 SLAVE 优先。Bit10 和 bit11 的区别是，bit11 的配置在强制情况下生效，PHY 只能工做与 bit11 指定的工作模式，而 bit10 的配置在非强制配置情况下生效，仅仅控制 PHY 在 AN 时候的优先级，偏向于 Maser 或者 Slave，但是最终的工作模式看协商的结果，不一定和优先级配置的结果一致。

1000BASE-T Full Duplex/ Half Duplex: 1000BASE-T 自动协商配置信息。在 1.4 节中曾经说明，寄存器 4 的自动协商通告信息寄存器没有关于 1000BASE-T 的信息，1000BASE-T 的自动协商通告信息由这两位进行配置，分别对应全双工和半双工两种工作模式。需要注意的是，1000BASE-T 工作模式的自动协商是强制的，也就是要想端口 1000BASE-T 模式工作正常，自动协商是必须 Enable 的。否则端口可能出现异常。

思考：一个千兆端口，其寄存器 0 读到的值为 0x0140，请问该配置是正确的还是错误的？为什么？

1.9 MASTER-SLAVE Status Register

寄存器 10 是 1000BASE-T 模式的状态寄存器，指示 PHY 及其 Linkpartner 的状态信息。详细的状态描述见表 7，表格中各个状态位的具体含义说明的相当清楚了，无需赘述。需要注意的是，关于 Linkpartner 的信息是通过自动协商完成的，而 1000BASE-T 的协商信息是通过 Next Page 交互的，因此只有在寄存器 6 中确认 Next Page 已经收到，寄存器 10 的 Linkpartner 信息才是有效的。否则有可能是错误信息。

表 7 MASTER-SLAVE Status Register

Bit	Name	Description	Type
10.15	MASTER-SLAVE configuration fault	Configuration fault, as well as the criteria and method of fault detection, is PHY specific.	RO/LH/SC
10.14	MASTER-SLAVE configuration resolution	1 = Local PHY configuration resolved to MASTER 0 = Local PHY configuration resolved to SLAVE	RO
10.13	Local Receiver Status	1 = Local Receiver OK 0 = Local Receiver not OK	RO
10.12	Remote Receiver Status	1 = Remote Receiver OK 0 = Remote Receiver not OK	RO
10.11	LP 1000T FD	1 = Link Partner is capable of 1000BASE-T full duplex 0 = Link Partner is not capable of 1000BASE-T full duplex	RO
10.10	LP 1000T HD	1 = Link Partner is capable of 1000BASE-T half duplex 0 = Link Partner is not capable of 1000BASE-T half duplex	RO
10.9:8	Reserved	Reserved	RO
10.7:0	Idle Error Count	Bits 10.7:0 indicate the Idle Error count, where 10.7 is the most significant bit.	RO/SC

Local/ Remote Receiver Status: 互连双方的 PHY 收发器状态信息。在

1000BASE-T 互联问题的故障诊断中，这是一个比较重要的定位信息，通过这个指示位，可以分别察看本地 PHY 和对端的 PHY 收发器是否正常，从而判断出问题出在哪一方身上。

Idle Error Count: Idle 错误计数器。1000BASE-T Link up 之后，其 MDI 信号不会有空闲状态。在没有数据帧发送的时候 PHY 会发送 Idle 信号。理论上说 Idle 信号的传输和数据信号的传输是一样的，如果 Idle 出错则数据往往也会出错，导致收发数据帧中出现 CRC。而在出现 CRC 的时候我们可以通过 Idle 计数器是否有错来初步判断出错的原因是，如果 Idle 也有错误，则说明原因可能与 MDI 相关，如果 Idle 没有错误，则原因可能在 PCS 以上的部分，或者是 MAC 的问题（当然这个判断不是绝对的）。不过需要注意的是，这个计数器是相当“脆弱”的，插拔网线都有可能导致 Idle 错误，因此在使用该计数器进行判断之前要先保证连接稳定，事先读一次寄存器 10 让 PHY 把计数器自动清零。

1.10 Extended Status Register

寄存器 15 是由 PHY 厂商在 PHY 中写入的指示 PHY 功能的状态寄存器，标明 PHY 是否具有 1000BASE-X 或者 1000BASE-T 的能力，实际应用和调试中实用价值不大。

Bit(s)	Name	Description	R/W ^a
15.15	1000BASE-X Full Duplex	1 = PHY able to perform full duplex 1000BASE-X 0 = PHY not able to perform full duplex 1000BASE-X	RO
15.14	1000BASE-X Half Duplex	1 = PHY able to perform half duplex 1000BASE-X 0 = PHY not able to perform half duplex 1000BASE-X	RO
15.13	1000BASE-T Full Duplex	1 = PHY able to perform full duplex 1000BASE-T 0 = PHY not able to perform full duplex 1000BASE-T	RO
15.12	1000BASE-T Half Duplex	1 = PHY able to perform half duplex 1000BASE-T 0 = PHY not able to perform half duplex 1000BASE-T	RO
15.11:0	Reserved	ignore when read	RO

2、PHY 扩展寄存器分析

除了 IEEE802.3 定义的 Register0-15 外，Register16-31 由 PHY 制造商自行定义，还有制造商通过分页存储技术扩展的更多寄存器空间，在这些寄存器中制造商定义了很多 PHY 的功能的控制以及状态指示信息，这些信息对我们在 PHY 的应用以及故障诊断中有的时候可以起到决定性的作用，但是由于这写寄存器不是 IEEE802.3 标准定义的，因此寄存器的地址以及功能名称在不同厂家的资料中有很大的差异，甚至在同一厂家的不同芯片中也不尽相同，因此下面的讨论只能就某一类的功能应用或者状态指示进行说明，但是其详细的名称和寄存器的地址要结合具体芯片具体分析，这里不能给出一个确切的答案。

目前主流的 PHY 都通过分页技术对 PHY 寄存器空间进行扩展，提供更多的寄存器空间来控制 PHY 更多的功能行为和提供更多的 PHY 状态指示信息。下

面针对我司惯常使用的 BCM 和 Marvell 的 PHY 寄存器分页存储技术进行简单描述，至于分页扩展出来的寄存器空间其具体作用千变万化，在此不作详细讨论。

BCM 的 PHY 绝大多数的寄存器地址空间都是单页地址，通过分页扩展的寄存器集中在 0x18 和 0x1C 两个地址上，0x18 和 0x1C 这两个寄存器专门定义了几个 bit 作为页地址 (Shadow Value)，其他的 bit 则是功能位，在不同的 Shadow Value 这些功能位将代表不同的功能。对 0x18 和 0x1C 的读写操作需要先修改 Shadow Value，然后才能访问到正确的寄存器空间。以下以 BCM5488S 为例，分别说明寄存器 0x18 和 0x1C 的访问方法。

寄存器 0x18 寄存器的 bit2: 0 定义为 Shadow Value，在需要读寄存器 0x18 的某个 Shadow 时，先要做一个写操作来切换 Shadow，这个写操作必须指定 bit15=0，bit14: 12 等于需要访问的 Shadow Value，bit2: 0=111，其它 bit 忽略，这时候 Shadow 切换成功；然后再对寄存器 0x18 进行读操作即可读到对应 Shadow 的寄存器值。在进行写操作的时候，则可以直接将 bit15: 3 等于需要写入的数据，bit2: 0 等于需要写入的 Shadow Value 即可完成需要的写操作（对 Shadow 111 的寄存器写操作有例外要求就是 bit15=1）。

寄存器 0x1C 的 bit14: 10 定义为 Shadow Value，bit15 定义为写使能位。在需要读寄存器 0x1C 的某个 Shadow 时，先要做一个写操作来切换 Shadow，这个写操作必须指定 bit15=0，bit14: 10 等于需要访问的 Shadow Value，其他 bit 忽略，即可切换 Shadow 成功，然后再对寄存器 0x1C 进行读操作即可读到对应 Shadow 的寄存器值。而如果需要写某个 Shadow 寄存器，则指定 bit15=1，bit14: 10 等于需要写的 Shadow Value，其他 bit 也置需要写的值，写入寄存器 0x1C 即可完成。

Marvell 的分页方式和 BCM 有很大的不同，Marvell 一般指定寄存器 0x16（寄存器 0x0-0x1c 的页地址）和寄存器 0x1D（寄存器 0x1E 和 0x1F 的页地址）作为页寄存器，几乎针对每个寄存器都有分页空间，因此在访问每个寄存器之前都必须弄清楚该寄存器的页地址，需要将页地址写入页寄存器中，然后再访问对应地址的寄存器即可。

注意：值得注意的是，软件的定期 Linkscan 操作常常会“暗中”修改 Shadow Value，因此在实际操作中如果发现切换 Shadow Value 不成功，要关闭 Linkscan 再尝试一下。在端口异常问题调试的时候，我们经常通过寄存器比较来查找问题的线索，然而正常的软件读命令只会读取默认 Shadow 的寄存器值，而很多隐藏在其他 Shadow 中的信息经常被我们忽略了，这个小小的忽略有时候会让你的调试查证工作走很大的弯路，切记！

2.1 工作模式控制器

现在的 PHY 芯片对外接口形式越来越丰富，在 MII 端可能有 GMII/RGMII/SGMII 等，在 MDI 端则可能有电口、光口等；这些不同的接口使得 PHY 可以支持各种不同的工作模式。但是在具体的应用中环境，PHY 只有设置为设计需要的工作模式才能够正常工作。因此，当遭遇端口不能 Link 之类的问题的时候，如果外围的电源时钟复位没有检查出什么异常的话，查看一下工作模式配置有没有正确，这是 PHY 能够正常工作的一个必要条件。

2.2 端口驱动模式

这是一个在 1000BASE-T 模式下需要关心的问题,1000BASE-T 需要在 Cat5 UTP 上传输 1000Mbps 的数据,其信道编码采用了 PAM-5 编码格式,MDI 上的信号为 5 阶电平格式,而其信号幅度和 100M 模式下一样是 $\pm 1V$,相应的其噪声容限就降低了,尤其在在百米长线连接的时候容易出现 Link 不稳定、CRC 错误等问题。1000BASE-T PHY 的 MDI 信号一般有 Class A 和 Class B 两种驱动模式,Class A 的驱动能力比 Class B 要强,因此在百米长线连接时候会表现出更好的性能。但是由于 Class A 的功耗比 Class B 要大,因此 PHY 的默认配置一般都是 Class B 模式。1000BASE-T 端口如果短线连接 ok 而长线连接有问题的时候可以尝试调整其端口驱动模式为 Class A 看看性能是否有改善。需要注意的是,有些 PHY 的这个控制位是隐藏的,在 DS 中可能差不到对应的配置位,因此就需要仔细阅读勘误表等资料或者向制造商咨询,以获得相应的配置方法。

案例: S5750-48GT/4SFP 农行生产所使用的 PULSE2 \times 6RJ-45,长线大约只能支持到 80 米。S5750-48GT/4SFP 使用的是 BCM5488 PHY,后换到 Marvel 的 PHY88E1145 上故障依旧。当时由此判断是 PULSE2 \times 6RJ-45 品质变异,于是更换成 Bel 的三合一,生产没多久又出同样的长线问题!看来并不是简单的更换部品就能了事。

在 88E1145 上做实验的是后注意到一个现象,出长线故障的 88E1145 端口配置一下 Super-link 命令,马上就可以解决问题。这个命令的具体操作就是将 88E1145 的端口驱动模式由 Class B 改为 Class A (以前 88E1145 也有出过长线故障问题,这样可以解决)。可惜察看 BCM5488S 的数据手册并没看到关于 Class B 和 Class A 的描述,向 BCM 方面咨询,他们给了一个勘误表,根据说明修改了一下寄存器,问题也解决了。这个寄存器是 BCM 的保留寄存器,BCM 的解释也是修改端口驱动模式的。

2.3 预加重配置

在高速的接口信号传输的时候,由于集肤效应以及电磁辐射的影响,信号的各个分量受到不同程度的衰减(一般是越高频的分量受到的衰减越大),信号经过传输线到达接收端的波形和输出端输出的信号波形并不是“相似的”,因此在发送端需要预先对信号的高频分量进行加重,特别是在万兆口应用中,需要通过接插件或背板传输的 XAUI/HIGIGA 等接口一般都需要对芯片的发送预加重(pre-emphasis)的调整,根据不同的应用环境进行不同的预加重配置,使得接收端获得比较好的接收眼图。除了 PHY 芯片,提供 XAUI/HIGIGA 接口的 MAC 芯片也有对应的内置 PHY,每个端口一般都有对应的预加重配置寄存器,在调试中需要根据眼图的测试结果进行配置。

案例: 福建工程学院两台 S8606 通过万兆线路互联, Ping 对端存在约 1%的丢包,更换万兆板、线卡槽位无法解决, CLI 上显示接口不存在 CRC 错误。针对这个问题,派黄赞到现场分析,通过分析端口的报文计数器,分析应该是在 CM 板和线卡之间发生了丢包,在进一步的检查万兆口的配置,发现 BCM5676 的 HG 端口发送信号预加重没有按照芯片要求进行配置,导致 CM 板和线卡之间传输出错。修正预加重配置后问题解决。需要注意的是,BCM5676 是一个 MAC 芯片,提供 4 个 XAUI/HIGIGA 端口。但是这些端口都有内置的 PHY 单元,需要进行正确的配置来确保信号的可靠连接。

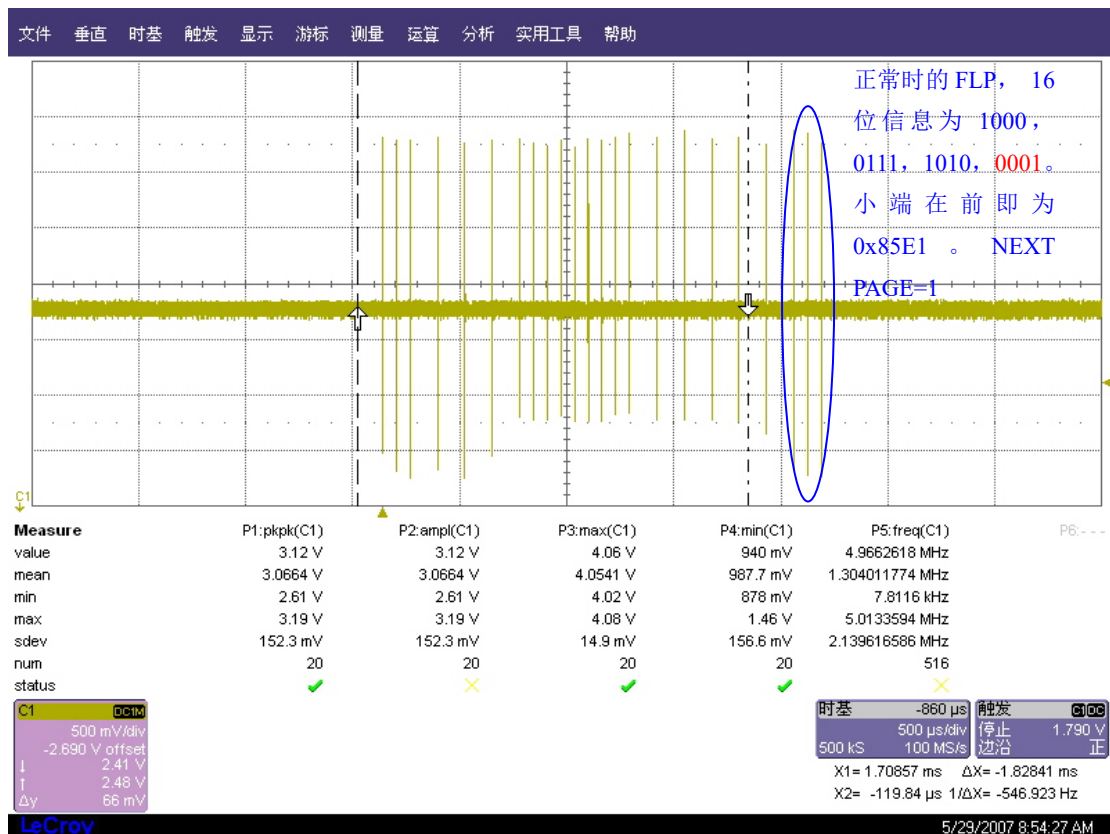
思考: 有些芯片除了发送预加重(pre-emphasis)的配置外,还有端口的接收均衡控制

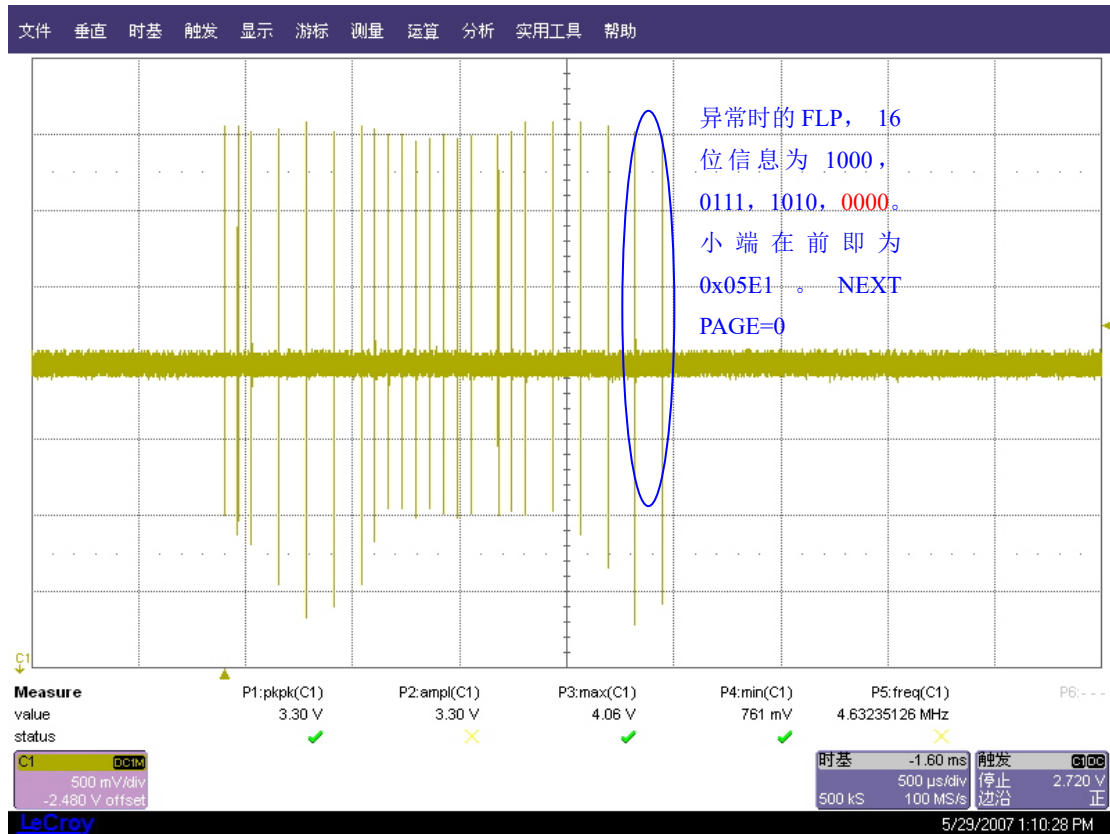
器 (Receive Equalization Control) (如 BCM56700 芯片), 这两个配置有什么区别? 分别起到了什么作用?

2.4 自动协商降格

由于 1000BASE-T 连接对 UTP 线缆的要求是比较苛刻的, 而 100M 连接对线缆的要求就相对宽松, 因此目前很多千兆电口的 PHY 芯片 (BCM 和 Marvell 都有, 典型代表 BCM5488S 和 88E1240, 不过 BCM5488S 这个功能名叫 Ethernet@WireSpeed mode。)提供了一个自动协商降格 (Downshift) 机制, 就是在 AN 阶段如果协商双方都支持 1000BASE-T, 但是协商完成后却不能建立 1000BASE-T Link, PHY 经过一定次数的尝试之后将自动降格到 100M 协商来建立 Link。这本来是个挺好用的功能, 但是由于不是 IEEE802.3 的标准, 因此芯片在实现上多少有些不够严谨。而且在实际应用中如果一个千兆口无端被协商成了 100M, 也容易对客户造成误解。因此该功能最好不要打开 (对应的寄存器 Downshift Enable 位置 0)。

案例: 还是 S5750-48GT/4SFP, BCM5488S, 调试中发现, 用 20 厘米短线将端口互联可以正常 Link, 不过调试中偶然发现将网线频繁的插拔几次, 最后接上以后发现端口居然 Link 状态为 100M, 当时百思不得其解。于是去测试端口的 FLP 信号, 发现正常时候和故障时候的 FLP 有差异如波形图所示。很显然 PHY 的自动协商发送信息自己没有通告 1000BASE-T 的能力, 自然不能 Link 到千兆。





再去仔细学习了一遍 BCM5488S 的相关资料，发现了一个 Ethernet@WireSpeed 功能，这个功能的描述为：When bit 4 = 1, Ethernet@WireSpeed mode is enabled. If the link cannot be established within two to nine attempts (the number of attempts is set by bits[4:2] in register 1Ch, shadow value 00100), the BCM5488S downgrades its advertised abilities and again tries to establish a link. When bit 4 is cleared, the BCM5488S advertises its abilities according to registers 04h and 09h。也就是说，当两个端口协商的努力失败后（协商次数取决于 5488S 的 0x1C 寄存器的 bit[4:2]，可在 2~9 次之间调整），5488S 会对其工作能力进行降级，即从千兆降到百兆。查到这里问题就比较清楚了，配置一下寄存器，关闭 Ethernet@WireSpeed 功能问题就解决了。

思考：在这个案例，发现自动协商信息和正常状态不一致是通过测试 FLP 看出来的，能不能通过分析寄存器来得到这些信息，需要看哪个寄存器，哪一端的寄存器？

2.5 Auto-Crossover 配置

所周知，两个以太网端口要正确地建立 Link，必须将一个端口的发送信号连接到另一个端口的接收端上（1000BASE-T 为同线双向，但是也要考虑线对顺序），否则的话会 Link 失败。以太网端口的 UTP 对外接口有两种模式，分别叫 MDI 和 MDIX，对应的，连接用网线也有直连线 and 交叉线之分。由于一个端口到底是 MDI 还是 MDIX 从外观上是看不出来的，为了免除人们调线连接的麻烦，IEEE802.3 后来定义了一个 Auto-Crossover 功能（还有个别名叫 Auto-MDIX），PHY 可以自动改变其 pin 脚为 MDI 或 MDIX 模式来建立正确的 Link。可是这个功能是要依赖 FLP 信号来实现的，因此在端口关闭 AN 的情况下 Auto-Crossover 功能也随之不起作用了。这是一个经常被忽略的问题，大家往往都认为可以 Auto-Crossover 的 PHY 可以随意的进行连接，在端口 Link 不上的时候潜意识中

就没有想过线连的对不对！

随着技术的进步，现在很多 PHY 开发出了不依赖 FLP 的 Auto-crossover 功能，可以在关闭 AN 的情况下依然有效，但是可惜的是这个功能一般有个控制寄存器，更可恶的是默认情况下是不打开的，因此在调试以及应用中要万分注意，看看 PHY 关于 Auto-crossover when AN is disable（当然也有可能是以其他词汇来描述的）的配置是不是配对了。要不然就相当于给自己留了一个陷阱。

案例：内蒙农行，我司 M8600-48GT 线卡和 cisco 7507 以太网口卡 PA-1FE-TX 互联，只能协商出百兆半双工速率，若强制成百兆全双工则不能 link。而使用 Cisco 2950 强制百兆全双工可以 link。从现象看，cisco 7507 以太网口卡 PA-1FE-TX 应该是不具备自动协商能力，强制 100M Full 配置 Link 不上，有两种可能，一则 M8600-48GT 的强制配置没有关闭自协商，另一种可能就是 MDI 连接错误。和软件确认，M8600-48GT 强制配置中肯定已经关闭自协商，核对 BCM5488S 的寄存器配置，发现 Force Auto-MDIX Mode 这个配置位被设置成 0 = Auto-MDIX is disabled when auto-negotiation is disabled，导致不能 Link。

思考：

- 1、这个案例中我们为什么可以推论 cisco 7507 以太网口卡 PA-1FE-TX 不具备 AN 能力？
- 2、如果不知道寄存器配置，如何方便的验证两个端口不能 Link 是不是因为没有 Auto-crossover 功能导致的？

2.6 MDI 信号边沿速率调整

PHY 的 MDI 接口信号是直接连接线缆和其他设备互联的信号，信号的好坏直接影响连接的正确性和兼容性，因此此 MDI 信号需要根据 IEEE802.3 的指标进行测试。在实际应用中，受实际设计应用环境以及外围器件的影响，MDI 信号有可能信号的上升下降时间变差，有些 PHY 提供了调整信号边沿速率的寄存器，可以根据需要和测试结果对 MDI 信号的边沿速率（Edge Rate）进行调整，以达到满足测试指标的要求。

案例：S2928G 进行 100BASE-T 的测试，发现信号的上升下降时间超出测试指标要求，最大值达到了 5.138ns（指标为 3~5ns）。一般来说遇上这种情况我们都会一筹莫展，因为信号是芯片输出出来的，而外围的阻容设计 MDI 接口也是相当经典成熟的设计，没有多少可以调整的余地。可是仔细的去学习 PHY 芯片 BCM5498 的相关芯片，发现其 Errata 中有关于 100BASE-T 信号测试超标的描述，按照资料的解决方法，修改芯片的寄存器配置信号就达标了。通过这一案例我想我们至少可以得到两点教训，其一是做设计一定要仔细学习芯片的勘误表，否则很多厂家已经发现并提供解决方案的问题我们可能由于无知而投入很大的时间和精力去查，吃力不讨好；其二就是，对于芯片输出的信号，我们并不是无可奈何的！芯片的输出信号受到很多因素的影响，芯片的工作电源、时钟还有内部的寄存器配置都可以改变芯片输出信号的指标，在这一点上我们应该放开思路，而不要因为是芯片内部的事情就望而生畏。

2.7 错误指示寄存器

在前面章节中我们有讨论过，PHY 的标准寄存器集有一些错误状态指示信息，比如 MASTER-SLAVE Status Register（寄存器 10）中有 Local Receiver Status，

Remote Receiver Status, Idle Error Count 等错误信息, 充分利用这些信息对我们在调试以及故障处理中定位问题原因将有很大的帮助。除了标准寄存器集定义的这些错误指示信息之外, PHY 厂商往往还定义了一些其他的错误指示信息, 方便实用者对链路状态进行监测和分析, 下面对一些常见的错误信息及其应用进行分析。

CRC Error Count: PHY 的 CRC 错帧统计计数器。本来 CRC 校验是 MAC 的功能, 但是现在很多 PHY 也实现了 CRC 校验的功能, 不过这个计数器一般只有 8bit, 也就是只能统计最多 256 个错帧。在查证 CRC 错帧问题的时候, 除了察看 MAC 的 MIB 以外, 还可以辅以 PHY 的 CRC 错帧统计计数器, 通过综合这些信息来初步判断, 出错的位置在 PHY 以上或者 PHY 以下。而且, 结合 Idle Error Count 也可以进行判断, 一般如果是 MDI 信号问题导致的 CRC 会伴随有 Idle 错误, 如果是对方 MAC 发过来的帧内容有误, 则应该不会出现 Idle 错误。这些信息在故障诊断中可以发挥相当重要的作用。

案例: S2951XG 常温烤机测试发现, 将端口两两短接起来, 从外面引入一个包发现跑不起来, 起初以为是软件配置有广播风暴抑制的原因, 检查软件配置却没有发现异常, 后来读取端口 MIB 发现有大量的 CRC 错包。于是判断跑不起来的原因是引入的帧转发后全错掉丢完了, 可是为什么出 CRC 呢? 除了 MIB 信息, 读取 88E1145 的 Idle Error Count 和 CRC error count 发现 PHY 也检测到了 CRC 错误, 同时伴随有大量的 Idle Error 记录, 因此判断故障原因应该和 PHY 底层信号有关。于是详细核对 88E1145 底层有关的配置, 发现软件对 MDI 接口的输出信号幅度 (VOD) 进行了调整, 将 VOD 提高了 14%, 这个配置是之前为了解决 88E1145 D0 版本芯片长线性不过而做的调整, 可能已经不适用于 S2951XG 上的 88E1145E1 版芯片。将 VOD 配置修改到默认配置, 故障解决。

Transmit Error: 发送错误指示。Transmit Error 是 IEEE802.3 描述的一个错误状态, MAC 芯片在发送的时候如果出现了某种错误, 则通过 MII (GMII) 的 TX_ER 信号通知 PHY, 这时候 PHY 将在 MDI 上发送一个特殊的码组代表 Transmit Error 状态。因此 PHY 的 Transmit Error 指示寄存器一般来说并不是指本端错误, 而是指 PHY 接收到了 Linkpartner 的 Transmit Error 信息, 这时候说明, 出错的位置应该是在 Linkpartner 的 MAC 上。

案例: 我司 RSR50 主板固化的千兆口和 M7600-24GT 模块互连的时候, M76 的端口会收到一些 CRC 错包, 在客户现场更换 H3C 的交换机和 RSR50 互连也一样发现 H3C 的交换机端口有收到 CRC 错包。该现象在我们本地实验室重现不出来, 给查证工作带来很大的难度。

在初期一度怀疑是 RSR50 兼容性问题, 因此在我们这边按照规范重新进行了一轮全面的兼容性测试, 没有发现任何异常问题。然后对 RSR50CPU 连接 PHY 的 GMII 接口进行测试, 各信号指标均达标, 基本上排除时序临界导致出错的怀疑。在分析 M76 的端口计数没有什么发现的情况下, 从现场读取 M76 的 PHY 寄存器回来分析, 发现其中一个疑点是 PHY 的状态寄存器中有指示端口检测到 Transmit Error。注意, 这里的 Transmit Error 并不是意味着端口自己发送发现错误, 而是 LinkPartner 通过特殊的编码 (如 4B5B 编码中的/H/码组) 发送过来的地表示其发送操作出错的指示。这个操作应该是 MAC 发现错误以后主动通报的, 因此故障怀疑的重点应该从 PHY、接口、链路等这些模块转移到 MAC 层, 重点查 RSR50 的 CPU (集成端口 MAC 模块) 设计和配置, 最终通过修改 CPU 的 FIFO 配置解决了问题。

在这个故障的解决过程中, 还走了一些弯路。其实由于对 Transmit Error 的含义理解不够透彻, 当时我们看到这个错误只是以后并没有明确的就开始把重点集中在 CPU BCM1250 上, 当时的想法是为了弄清到底什么样的包出了 CRC, 于是寄了一台 S2924G 过去作为抓包

之用（S2924G 有个功能可以配置端口不进行 CRC 校验从而将 CRC 包原样转发出来而不过滤），但是却发现将 S2924G 连接到 RSR50 之后没有再出 CRC 错包了，经分析怀疑 S2924G 的 MAC 对 Transmit Error 的处理可能和 M76-24GT 不一样，于是将 RSR50 上 GMII 接口的 TX_ER 信号断开，再和 S2924G 连接，再观察也出现 CRC 错包了。到此才确定，出错的原因肯定是 BCM1250 内部出了问题。

这个案例给我们的启示是，对于芯片的错误指示信息，我们一定要弄清楚各项错误只是代表的含义和产生的原因，才能顺藤摸瓜，找到故障的本源。

除了以上列举出来的错误指示，PHY 其实还有很多错误指示寄存器可以为我们的故障定位分析提供帮助，比如 AN Error, Parallel Detect Error 等这些错误信息含义都一目了然，但是在端口故障的时候确实非常重要的证据信息。还有新的芯片可能还提供了其他的错误指示，其含义的作用需要结合芯片的相关资料 and 标准规范学习研究，在实际工作中加以应用。