Freescale Semiconductor
Application Note

Document Number: AN3869
Rev. 0, 06/2009

Implementing SGMII Interfaces on the PowerQUICC™ III

by *Networking and Multimedia Group*
Freescale Semiconductor, Inc.
Austin, TX

1 Introduction

SGMII is a serial interface for gigabit Ethernet that replaces previous standards like GMII and RGMII. This interface requires fewer physical pins, so it simplifies hardware routing and layout. However, beccause a SERDES function (internal to the PowerQUICC III™) is added within the chain of devices comprising the link, additional configuration (compared to other interface types) is required. This application note explains these configuration requirements to assist users to successfully program and set up this interface type. It also provides a brief description of SGMII in general and describes how it is similar and different to other gigabit Ethernet standards and interfaces. The specifics of the Freescale SGMII implementation are detailed, including steps and requirements to enable SGMII, configuration considerations and details of the steps required to successfully initialize a gigabit Ethernet port configured with SGMII as its interface. The intent of this application note is to provide the specifics required for any Freescale SGMII interface. Specifics to a particular Freescale part are only to illustrate the concept being described.

Contents

1. Introduction 1
2. Implementing SGMII Interfaces 2
3. Tips and Hints 12
4. Testing and Validation 12
5. Conclusion 12
6. References 13
7. Revision History 13

freescale™ semiconductor

1.1 SGMII and 1000Base-X Overview

SGMII is a non-IEEE-standard mode of communication (defined by Document ENG-46158 *Serial-GMII Specification* from Cisco Systems) between the MAC and PHY that allows for copper 10/100/1000BASE-T (IEEE Std 802.3ab™) operation. It replaces the classic 22-wire GMII connection with a low pin count, 2 or 4-pair, differential SGMII connection. SGMII defines a method for operating 10 Mbps and 100 Mbps MACs over the interface. In 100BASE-T mode, the MAC still transmits to the PHY at 1.25 Gb/sec, but each byte is repeated 10 times. The PHY then converts this repeated data to 100BASE-T format. The process is the same in 10BASE-T mode but each byte is repeated 100 times.

Where SGMII provides capabilities for three speed Ethernet over copper, 1000BASE-X is used in industry to refer to gigabit Ethernet transmission over fiber and identifies various Gigabit Ethernet physical layer standards (1000Base-LX, 1000Base-SX) as defined in IEEE 802.3z.

The SGMII specification is closely related to 1000Base-X in that each utilizes the same Physical Coding Sublayer (PCS) and Physical Medium Attachment (PMA) layers as is defined by 1000Base-X 802.3 Clause 36. Within the PQIII devices the PCS and PMA layers are implemented within the eTSEC via the Ten Bit Interface (TBI). The PMA interface operates at 1.25 Gbaud for both modes as is defined by IEEE 802.3 Clause 36. For SGMII this link is connected to an external Ethernet PHY device that supports copper based 10BASE-T, 100BASE-T, and 1000BASE-T operation. For 1000Base-X the PMA link connects directly to an optical transceiver (or across a backplane). Both SGMII and 1000Base-X PCS connect to the system side via the Ethernet MAC as is described in further detail in section 2.2.1 Data Link. What distinguishes SGMII from 1000Base-X is the information that is exchanged during the auto-negotiation sequence (Defined by IEEE 802.3 Clause 37) as well as speed resolution and rate adaptation that allows SGMII to operate at three speeds.

2 Implementing SGMII Interfaces

2.1 Hardware Configuration

The Freescale SGMII implementation leverages the internal SERDES (serial/deserial) block(s) within the chip. Refer to the applicable device reference manual to determine how many SERDES blocks are within the part. Also, take note of any SERDES blocks or lanes that may be shared with other interface types such as PCI-Express or serial Rapid IO.

When the SERDES lanes are operating as an SGMII interface, the clock recovery function is included within the interface. The advantage of this is that each interface only requires four signals (two pairs) for implementation (+TX_D, -TX_D, +RX_D, -RX_D), so there is no separate clock or clock routing required per interface. However, to make the clock recovery operates properly, a reference clock is required for each SERDES block (which may contain multiple SERDES lanes). In order to ease clock selection, a PLL multiplier is provided so the full interface clock speed is not required; the user may choose either a 100-MHz or 125-MHz clock source for the reference.

After the user determineS the SERDES block usage (SGMII, PCI-Ex, SRIO) and selects the reference clock frequency (100 or 125 MHz), the part must be configured to operate with these selections. This is done via signal strap options that are detailed in the power on reset configuration section of the applicable device reference manual.

NOTE

Because the hardware configuration is determined at POR (power on reset) the SERDES reference clock must be present when the reset signal is deasserted. Failure to do so can cause the reset sequence to hang as the hardware interface blocks wait for the SERDES PLLs to lock.

As an example, consider an MPC8544-based design with the following:

- 2 PCI-Ex ports
- 1 SGMII port on eTSEC1
- 1 RGMII port on eTSEC3
- 100 MHz SERDES reference clock

Table 1 lists the configuration signals that must be set (refer to Section 4.4.3 of the *MPC8544E PowerQUICC™ III Integrated Host Processor Family Reference Manual*).

Table 1. Configuration Signal Settings

Configuration Pin Name	Setting	Comments
cfg_IO_ports[0:2]	0b101	Enables SGMII usage in the SERDES, defines 2 PCI-Ex ports. Note that the reference manual defines the SERDES lane signal definitions for the PCI-Ex and SGMII interfaces.
cfg_tsec1_serial	0	Defines eTSEC1 as SGMII
cfg_tsec3_serial	1	Defines eTSEC3 as not SGMII
cfg_tsec1_reduce	1	Leave as default
cfg_tsec3_reduce	0	Defines eTSEC3 as a reduced interface type interface, used in combination with cfg_tsec3_prtcl to define RGMII
cfg_tsec1_prtcl[0:1]	0b11	Leave as default
cfg_tsec3_prtcl[0:1]	0b10	Defines eTSEC3 as GMII type interface, used in combination with cfg_tsec3_reduce to define RGMII
cfg_srds_sgmii_ref_clk	1	Defines SERDES reference clock as 100 MHz

This is a specific example for the MPC8544 to illustrate the types of configuration signals expected, the signal definitions and settings are likely slightly different for other products. See the configuration signal section of the applicable device reference manual.

2.2 Link Overview

2.2.1 Data Link

Because there are many individual functional elements within the SGMII data link, it is important to understand how they interrelate.Within the Freescale implementation, the eTSEC MAC is configured as a TBI interface and internally connected to the appropriate SERDES interface. The internal connection to the TBI and SERDES blocks run at a rate equal to 1 Gbps, regardless of the external link speed, which may be 10 Mbps, 100 Mbps, or 1 Gbps. This is implemented by a rate block that repeats the characters as discussed previously in Section 1.1, “SGMII and 1000Base-X Overview.” See Figure 1 for specifics on

how these functional elements work together. Note that because the TBI block is utilized, the user must configure and initialize that as well as specify that the SERDES block is used.

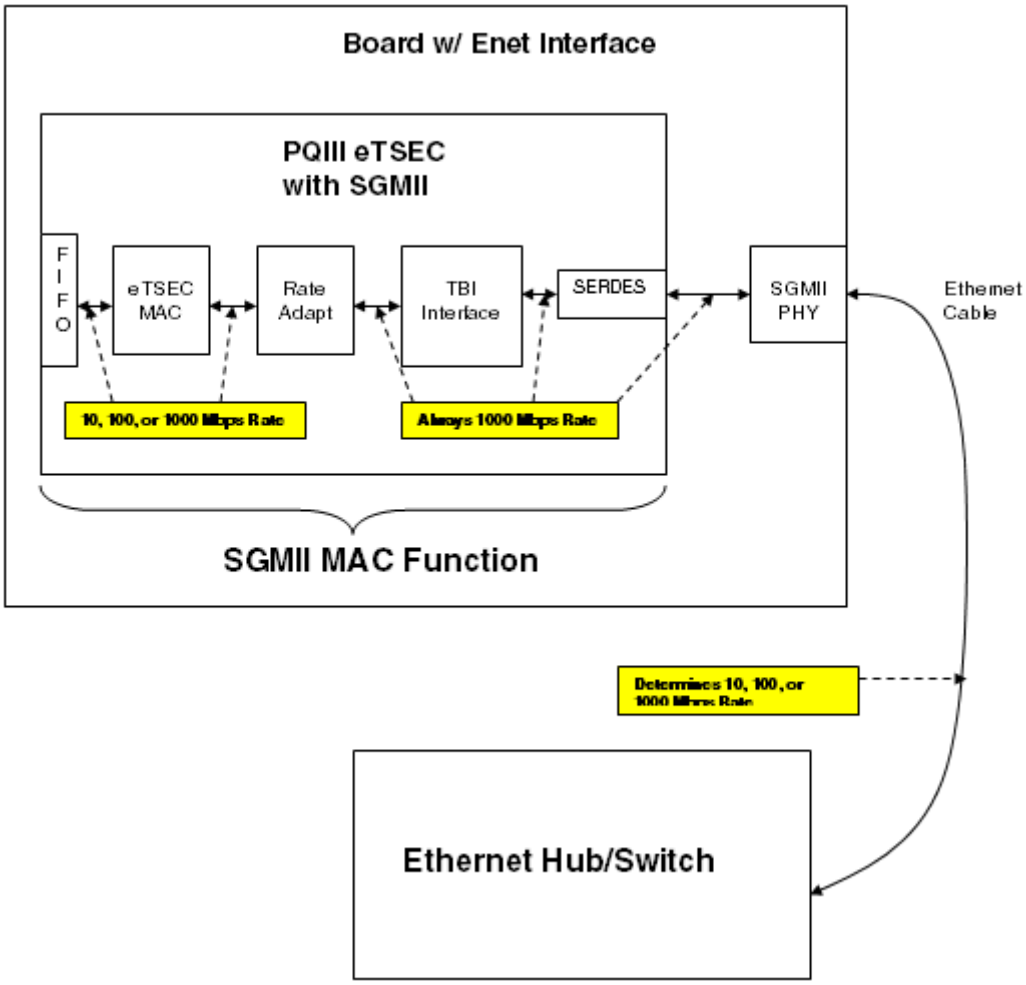

Figure 1. Functional Elements of the SGMII Link

2.2.2 Control Link

In a user system with SGMII there are multiple elements that need to be configured and initialized; in effect, there is a TBI interface with an internal SERDES phy that connects to an external SERDES phy. From a control standpoint this has more elements to consider than other eTSEC interfaces such as GMII or RGMII. Typically the MDIO management interface is used to configure the external phys; however, this same facility must be used to configure the internal TBI block/SERDES combination previously discussed. Depending on the PowerQUICC part used, some eTSEC's may or may not provide an external MDIO management interface. However, the TBI block for each eTSEC is associated in a one to one manner, even if that eTSEC does not have an external MDIO connection.

Figure 2 shows an example system configuration for the MPC8572.

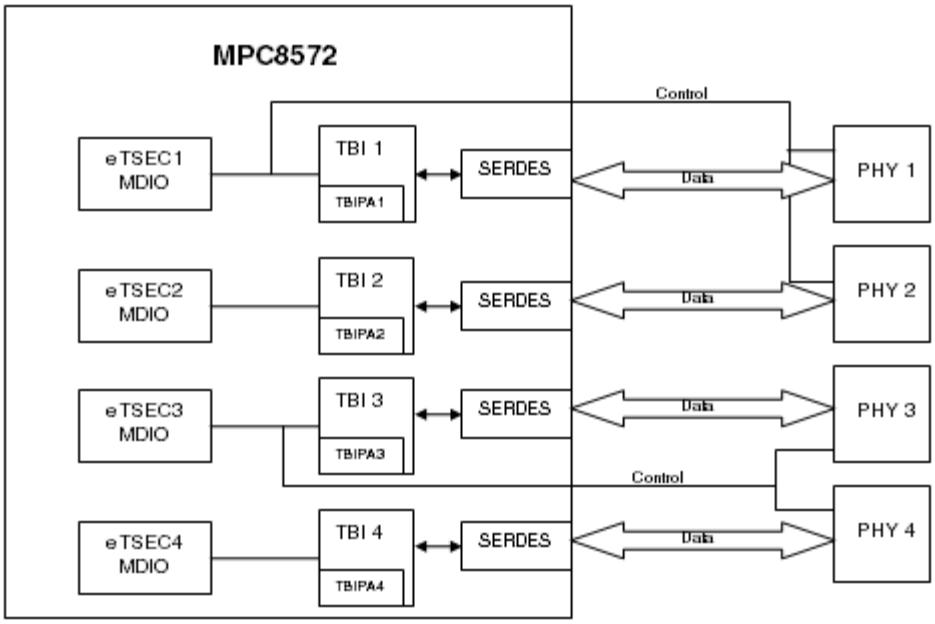

Figure 2. Controlling SGMII Functional Elements

Note that there are likely multiple phy "elements" controlled by the eTSEC controllers with external MDIO interfaces; this is required so that the external phy devices can be configured. The implication is that it is very important for the user to ensure that all the phy elements on the same MDIO interface have different phy addresses. For the internal TBI block this is controlled by the TBIPA register for each eTSEC block. The reset value of this register is 0x0, which is not a valid phy address. Therefore this register *must* be initialized for each TBI (thus SGMII) port in the system. For external phy devices the address is typically a pin strapping option, so the designer must ensure that the phy addresses of the external phys are different from any internal TBI that may be sharing that management interface.

Also note that for the TBI block the user must use the management interface directly associated with the eTSEC, however for external phy configuration this may not always be the case. For example, for the system configuration above, eTSEC1's management interface is used to configure and manage the TBI for eTSEC1 and the external phys connected to eTSEC1 and eTSEC2

2.3 Auto-Negotiation

Auto-negotiation allows devices based on several Ethernet standards, from 10BaseT to 1000BaseT, to coexist in the network and reduce the risks of network disruption that could arise from incompatible technologies. This capability helps ensure a smooth migration path from Ethernet to Fast Ethernet and Gigabit Ethernet. Auto-negotiation is the communication or handshake between two remote devices to choose common transmission parameters, such as speed and duplex mode. The connected devices first share their capabilities for these parameters and then choose the fastest transmission mode they both support.

The following sections focus on how both SGMII and 1000BaseX auto-negotiation is accomplished using Freescale Power QUICC III devices.

2.3.1 SGMII Auto-Negotiation

Figure 3 illustrates SGMII configuration and auto-negotiation. Auto-negotiation is controlled and monitored via the PQIII ten-bit interface (TBI) and the TBI MII set of registers.

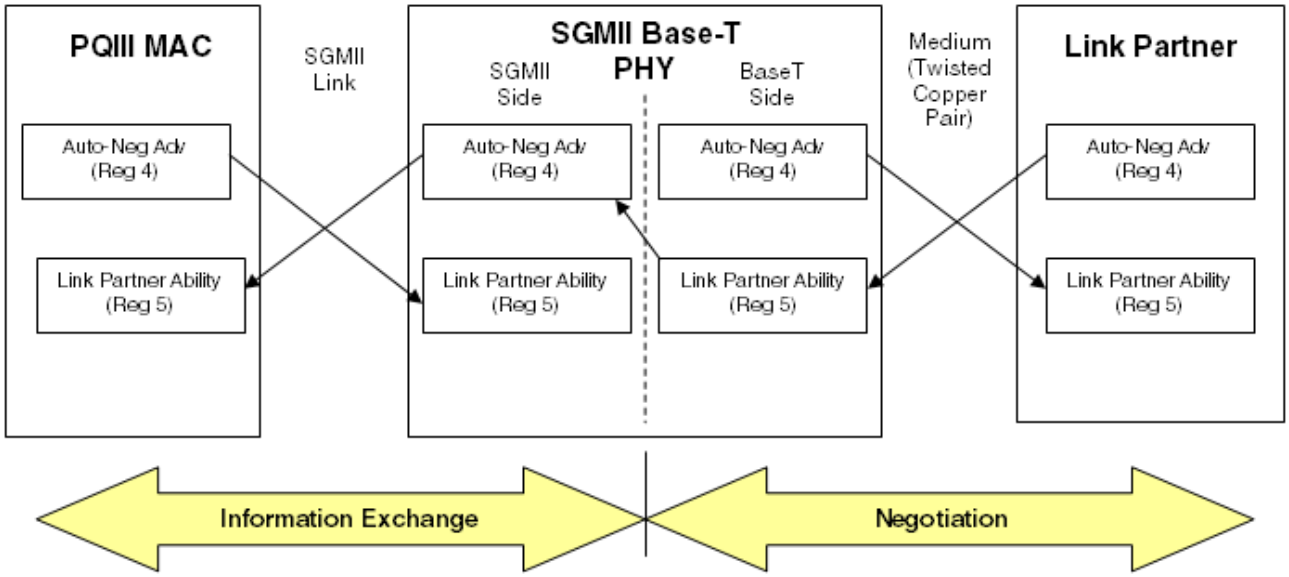

Figure 3. SGMII Link Exchange

SGMII is the mode of communication between the MAC and PHY that allows for copper 10/100/1000BASE-T (IEEE 802.3ab) operation. The speed and duplex of the copper link is agreed upon by link partners who advertise each other's capabilities using auto-negotiation, which starts automatically when any of the following conditions are met:

- Power-up/reset
- Loss of synchronization
- The link partner initiates auto-negotiation
- An auto-negotiation restart is requested—see bit 6 of the control register

Figure 4 shows the auto-negotiation control register within the PQ3 TBI register set. The default value of this 16-bit register is 0x1140, which enables auto-negotiation by default via bit 3. The default setting of the CR should be left as is for proper AN operation.

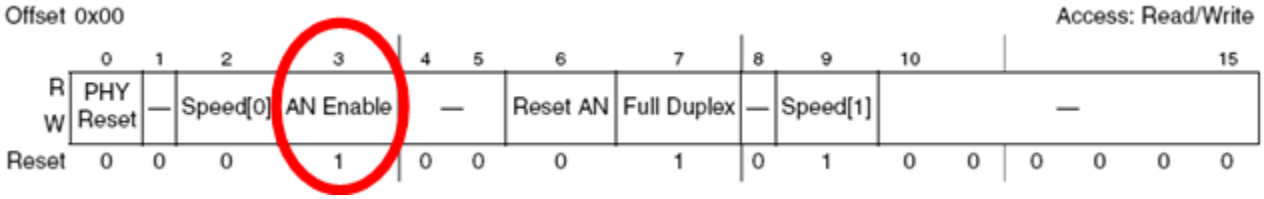

Figure 4. Control Register (CR) Format

There are two levels of auto-negotiation. The PHY auto-negotiates parameters such as speed and duplex mode with its link partner over the copper link per IEEE 802.3 Clause 27. The link partner's capabilities

are then transferred by the PHY to the PQIII MAC over the SGMII link using the auto-negotiation functionality defined in IEEE 802.3z Clause 37.

Parameters advertised by the PHY to its link partner are based on configuration of the PHY auto-negotiation advertisement register (register 4). The PHY internal registers are accessed using the MII management registers and interface of the PQIII processor.

The capabilities advertised by the remote partner during auto-negotiation are transferred into the PHY's link partner ability base register (register 5). Under normal conditions, this completes the Auto Negotiation information exchange.

The host processor can determine when the PHY auto-negotiation has completed by polling the PHY's auto-negotiation completion bit (bit 5) of the status register (register 1). Alternatively, the PHY can issue an interrupt to the host processor (PQIII) signaling the completion of the auto-negotiation cycle.

The auto-negotiation results are then passed to the PQIII MAC from the PHY using Clause 37 auto-negotiation. As is specified by the Cisco SGMII Specification, instead of the ability advertisement, the PHY sends control information during the Clause 37 auto-negotiation. ***While this exchange is labeled "Auto-Negotiation," it is really just an information exchange where the PHY passes the copper side AN results to the MAC.*** The MAC simply responds back to the PHY with and acknowledgement of receiving the results. Figure 5 shows the control information sent to the MAC from the PHY.

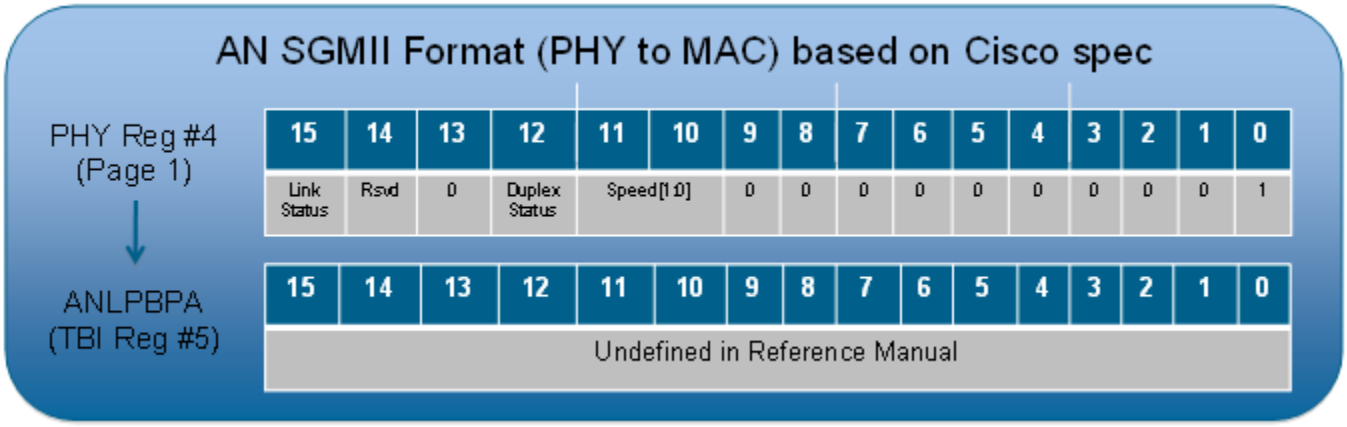

Figure 5. Phy to MAC (PQIII) SGMII Information Transfer

The MAC acknowledges the update of the control information received from the PHY and responds by asserting bit 14 of the TBI AN Advertisement Register (ANA, register 4). This is shown in Figure 6.

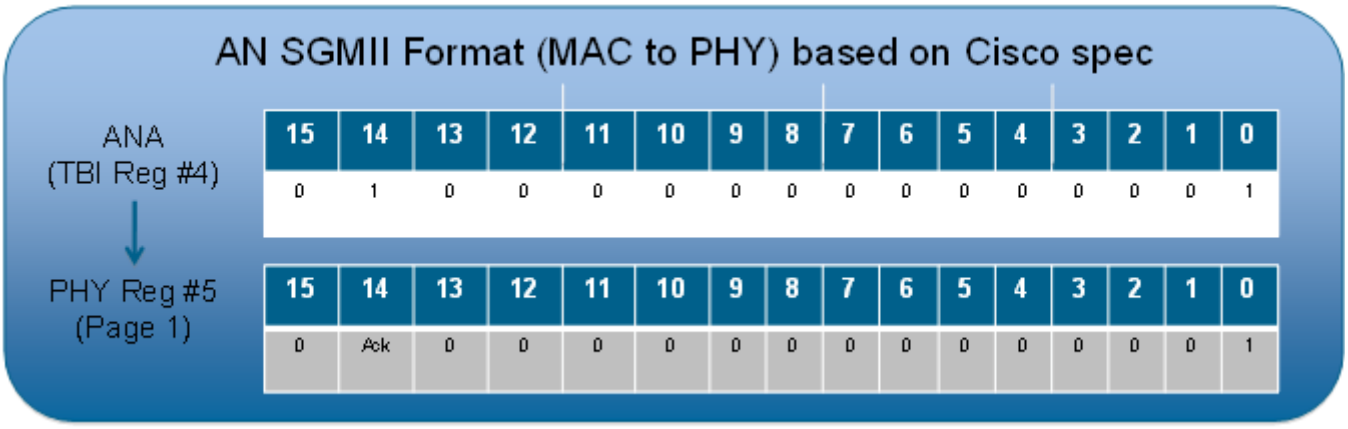

Figure 6. MAC (PQIII) to Phy SGMII Information Transfer

Upon completion of the control information exchange between the MAC and PHY, the auto-negotiation status register, shown in Figure 7, should read a value of 0x016D indicating auto-negotiation complete and valid link status.

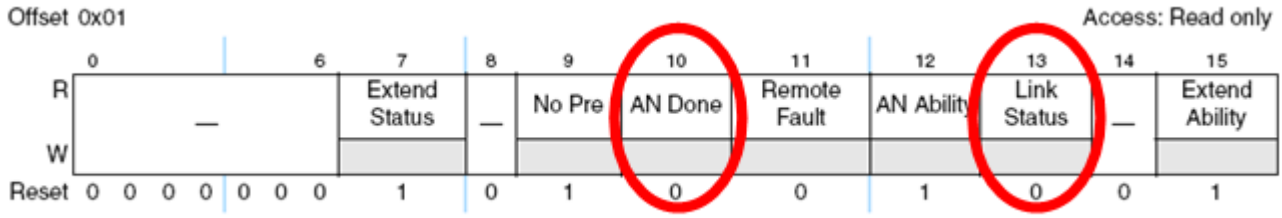

Figure 7. Status Register (SR) Format

Note that Extend Status (bit 7), No Pre (bit 9), AN Ability (bit 12) and Extend Ability (bit 15) are all read-only bits and return 1 when read. These bits do not reflect control information relayed by the PHY but rather just the ability of the PQIII. For example, AN ability is just the ability of the device and it is simply indicating this device can perform auto-negotiation.

2.3.2 TBI Configuration

SGMII auto-negotiation between the PQIII MAC and external PHY is accomplished through the TBI MII set of registers. Writing the TBI registers is completed in the same way as accessing an external phy by using the MII Management interface. The TBI registers are accessed as described in Section 2.2.2, “Control Link.”

2.3.3 1000BaseX Auto-Negotiation

1000BASE-X is used in industry to refer to gigabit Ethernet transmission over fiber and identifies various Gigabit Ethernet physical layer standards as defined in IEEE 802.3z. Figure 8 shows the 1000BaseX auto-negotiation.

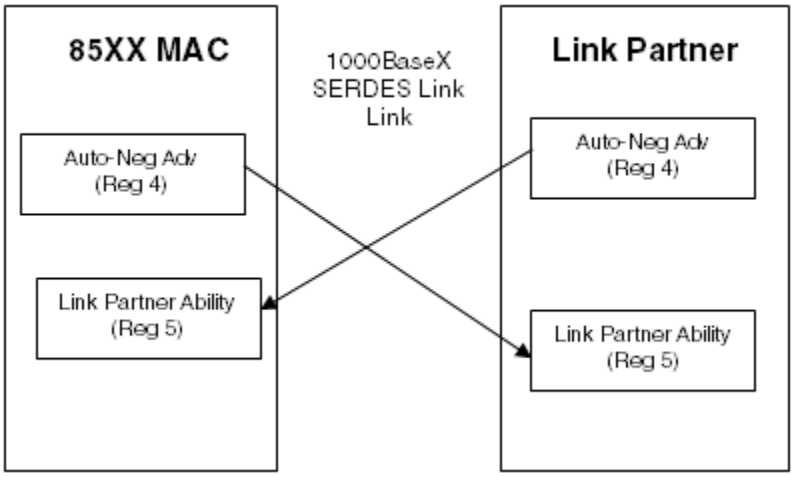

Figure 8. 1000BaseX Auto-Negotiation

1000BaseX auto-negotiation is similar to SGMII auto-negotiation as described in the above sections where both modes adhere to IEEE 802.3z Clause 37 auto-negotiation. What differ from SGMII are the capabilities that are advertised by the PQIII MAC as is specified by the PQIII TBI AN Advertisement Register (Register 4). Different from SGMII and based on a MAC to PHY information exchange, 1000BaseX is based on a MAC to MAC communication where capabilities are advertised to a link partner and the corresponding capabilities that the link partner advertises are detected by the MAC.

Configuration of the TBI registers (CR, TICON) remains consistent with what was described above for SGMII. For 1000BaseX auto-negotiation the MAC advertises duplex, Fault Condition signaling and Flow Control capabilities for the attached MAC to the remote link partner. Figure 9 shows the TBI 1000BaseX AN Advertisement register (register 4).

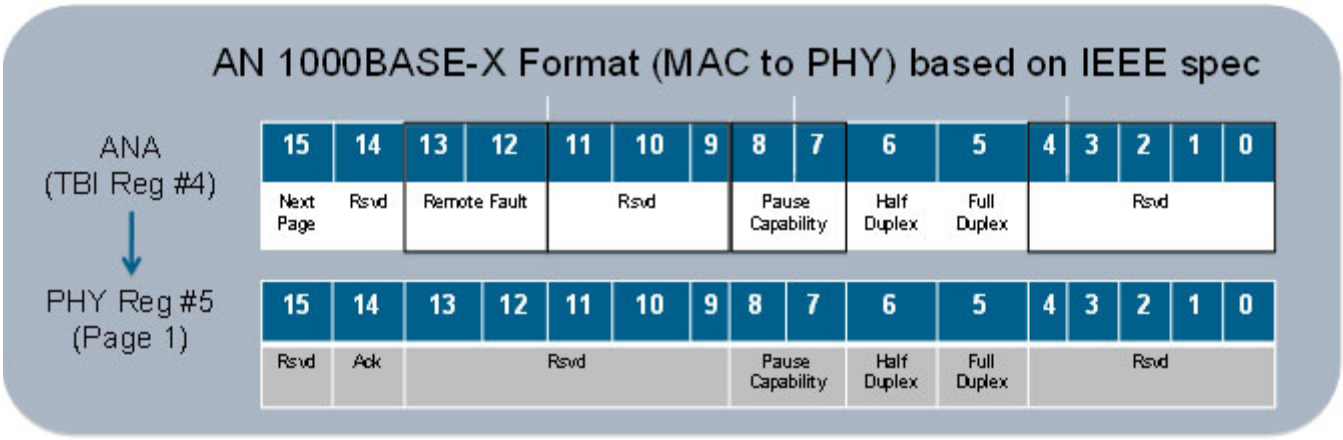

Figure 9. MAC to Phy 1000Base-X Information Transfer

The advertised abilities of the Link Partner are simultaneously transferred into the TBI auto-negotiation link partner ability base register (register 5), shown in Figure 10. This register contains the same information as in the auto-negotiation advertisement register.

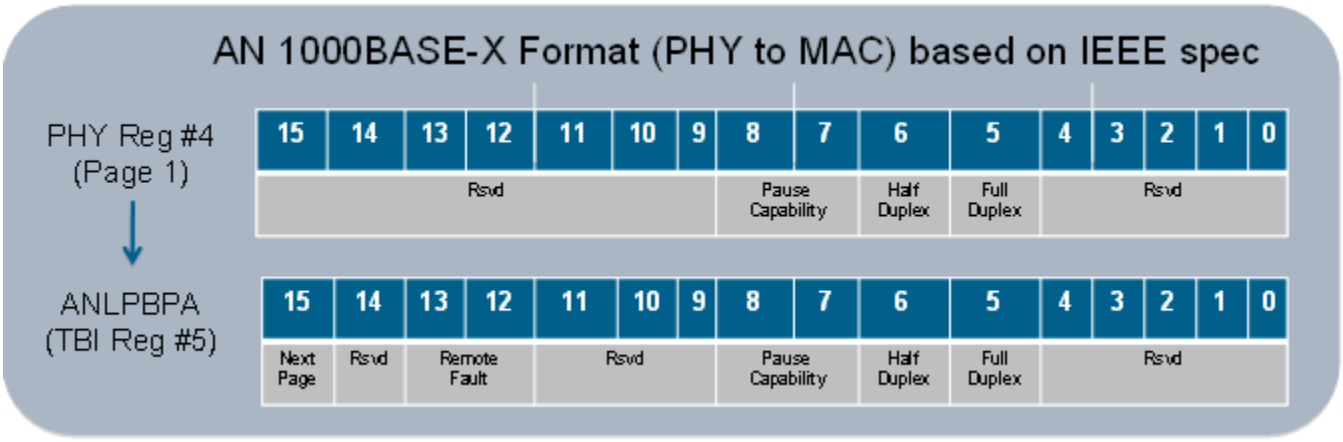

Figure 10. Phy to MAC 1000Base-X Information Transfer

2.4 Port Initialization Steps

Using the MPC8572 as an example the primary steps for configuration are as follows:

- POR configuration signal definition
- Initialization of the configuration registers mapped within the eTSEC controller
- Initialization of the TBI controller

Using the MPC8572 as a specific example, the eTSECs utilize the lanes of SERDES2. SERDES2 has its own programming options but the default values are used with respect to the writing of this application note.

POR Signals are as follows:

- LA[28] sampled low selects eTSEC1 for SGMII operation on SerDes2, lane 0.
- LGPL1/LFALE sampled low selects eTSEC2 for SGMII operation on SerDes2, lane 1.
- TSEC3_TXD[3] sampled low selects eTSEC3 for SGMII operation on SerDes2, lane 2.
- UART_SOUT[0] sampled low selects eTSEC4 for SGMII operation on SerDes2, lane 3.
- TSEC_1588_TRIG_OUT sampled low expects a 125MHz SerDes2 reference clock.
- TSEC_1588_TRIG_OUT sampled high expects a 100MHz SerDes2 reference clock.
- * No matter what SerDes2 clock reference is chosen, the SGMII interface operates at 1.25 Gbaud and supports 10/100/1000 Mbps data rates.

Memory-Mapped MAC Configuration Registers are as follows:

- 10Mbps:
  - ECNTRL[TBIM]=1, ECNTRL[SGMIIM]=1, ECNTRL[R100M]=0
  - MACCFG2[I/F Mode]=0b01
- 100Mbps:
  - ECNTRL[TBIM]=1, ECNTRL[SGMIIM]=1, ECNTRL[R100M]=1
  - MACCFG2[I/F Mode]=0b01
- 1Gbps:
  - ECNTRL[TBIM]=1, ECNTRL[SGMIIM]=1, ECNTRL[R100M]=0
  - MACCFG2[I/F Mode]=0b10

TBI Registers (accessed via memory-mapped registers MIIMCOM, MIIMADD, MIIMCON, and MIISTAT) are as follows:

- CR at offset 0x00
- SR at offset 0x01
- ANA at offset 0x04
- TBICON at offset 0x11

Configuration steps (MIISTAT contains the result of MII mgmt read cycles) are as follows:

1. Power-On-Reset (POR) signals select SGMII interface
2. Perform any desired SerDes2 programming.
3. Soft Reset the MAC (refer to procedure in the PowerQUICC reference manual)
4. Program TBIPA[TBIPA] accordingly (refer to section x.x.x). Choose a unique physical PHY address to ensure no conflict with any other external PHY.
5. Program the MII Mgmt Clock speed to <= 2.5Mhz via MIIMCFG.
6. Read TBI CR to ensure CR[AN Enable] is set, CR[Speed[0]]=0, CR[Speed[1]]=1, CR[Full Duplex]=1.

7. Program TBI TBICON[Clock Select]=1
8. Program TBI ANA = 0x4001 (SGMII) or 0x01A0 (1000BASE-X).
9. Set TBI CR[Reset AN].
10. Read TBI SR to ensure SR[AN Done]=1 and SR[Link Status]=1.
11. Read external PHY for speed and duplex information.
12. Program ECNRTL and MACCFG2 based on speed and duplex information.
13. Program all other MAC related registers as desired.
14. Enable Rx & Tx via MACCFG1.

3 Tips and Hints

There are checklist items to ensure proper operation based on the speed and duplex desired, as follows:

- The SerDes2 interface has two source clock speed options (100 MHz or 125 MHz). Verify the speed selected by the POR signal (TSEC_1588_TRIG_OUT) matches the actual source.
- ECNTRL[R100M] is only set for 100Mbps operation. Verify this is cleared for 10Mbps and 1 Gbps operation.
- TBI register field TBICON[Clock Select] at offset 0x11 must be set to 1.
- Timestamping is not supported for 10/100 Mbps modes. Verify RCTRL[TS]=0.
- The SGMII PHY being used may be sensitive to what is programmed in TBI register ANA at offset 0x4. Program accordingly based on the PHY's mode (SGMII or 1000BASE-X).
- Consult the latest published errata for any additional precautions. In addition to part errata, be sure to also check any documentation errata to ensure documentation errors are not propagated to your design.

4 Testing and Validation

The concepts discussed in this application note have been found to be critical for successful operation in a number of Freescale and customer systems. However, since there are limitless system factors (external phys, MAC to MAC operation…) possible with SGMII some modification and may be necessary for the specific application.

5 Conclusion

SGMII is an increasingly popular Ethernet interface due to the reduced number of signals required by the interface. However, compared to other Ethernet interfaces (GMII, RGMII) there are additional considerations that must be taken into account in order to successfully use this interface. These considerations are driven by the fact that each end-to-end link actually has additional phy or "phy-like" elements that impact the configuration and initialization steps required as well as alters functional considerations such as the true meaning of "auto-negotiation" in this environment. The points discussed in this application note assist the system designer and programmer by highlighting the key points required for successful operation.

6 References

- Cisco Systems Document ENG-46158 *Serial-GMII Specification*
- MPC8572E PowerQUICC III Integrated Host Processor Reference Manual
- MPC8544E PowerQUICC III Integrated Host Processor Reference Manual

7 Revision History

Table 2 provides a revision history for this application note.

Table 2. Document Revision History

Rev. Number	Date	Substantive Change(s)
0	06/2009	Initial public release.

How to Reach Us:

Home Page:
www.freescale.com

Web Support:
http://www.freescale.com/support

USA/Europe or Locations Not Listed:
Freescale Semiconductor, Inc.
Technical Information Center, EL516
2100 East Elliot Road
Tempe, Arizona 85284
1-800-521-6274 or
+1-480-768-2130
www.freescale.com/support

Europe, Middle East, and Africa:
Freescale Halbleiter Deutschland GmbH
Technical Information Center
Schatzbogen 7
81829 Muenchen, Germany
+44 1296 380 456 (English)
+46 8 52200080 (English)
+49 89 92103 559 (German)
+33 1 69 35 48 48 (French)
www.freescale.com/support

Japan:
Freescale Semiconductor Japan Ltd.
Headquarters
ARCO Tower 15F
1-8-1, Shimo-Meguro, Meguro-ku
Tokyo 153-0064
Japan
0120 191014 or
+81 3 5437 9125
support.japan@freescale.com

Asia/Pacific:
Freescale Semiconductor China Ltd.
Exchange Building 23F
No. 118 Jianguo Road
Chaoyang District
Beijing 100022
China
+86 10 5879 8000
support.asia@freescale.com

For Literature Requests Only:
Freescale Semiconductor
Literature Distribution Center
1-800 441-2447 or
+1-303-675-2140
Fax: +1-303-675-2150
LDCForFreescaleSemiconductor
@hibbertgroup.com

Document Number: AN3869
Rev. 0
06/2009