A Training on High-Speed I/O Interfaces: **SATA 3**

FTF-NET-F0155

Zhongcai Zhou | Application Engineer

A P R . 2 0 1 4
# History of SATA

## Pre-history:
- Parallel ATA (IDE, EIDE, …)
  - Higher pin counts
  - Wide cable
  - Low performance (Up to 133 Mbytes/sec)

## Speed

<table>
<thead>
<tr>
<th></th>
<th>SATA 1</th>
<th>SATA 2</th>
<th>SATA 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>1.5 Gbps</td>
<td>3.0 Gbps</td>
<td>6.0 Gbps</td>
</tr>
<tr>
<td></td>
<td>150 Mbytes/sec</td>
<td>300 Mbytes/sec</td>
<td>600 Mbytes/sec</td>
</tr>
<tr>
<td>Introduction time</td>
<td>2003</td>
<td>2005</td>
<td>2009</td>
</tr>
<tr>
<td>Spec revision</td>
<td>1.0a</td>
<td>2.5, 2.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Major New Features</td>
<td>baseline</td>
<td>- Native command queue</td>
<td>- Automatic Partial to Slumber Transitions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Port Multiplier</td>
<td>- Serial ATA NCQ Streaming Command</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Port Selector</td>
<td></td>
</tr>
</tbody>
</table>
SATA 3 and new features

- The full 3.0 standard was released on May 27, 2009
- Considering 8b/10b encoding, the maximum data rate:
  \[6.0 \text{ Gb/s} \times 0.8 = 4.8 \text{ Gb/s} \] (600 MB/s)
- Backward compatibility with SATA 3Gb/s connectors and cabling
- Enhanced power management capabilities
- Advanced NCQ (Native Command Queuing) streaming
- Two new form factor connectors
  - A small low insertion force (LIF) connector for more compact 1.8-inch storage devices
  - A connector designed to accommodate 7 mm optical disk drives for thinner and lighter notebooks
Cabling Recommendation

- Cable and connector is backward compatible with SATA rev2.6. This helps to smooth the migration from SATA 3.0Gbps to SATA 6.0Gbps.

- In the transition to SATA 6Gb/s, it is important to use high-quality cabling.

- Cables made from marginal materials that perform at the edges of SATA 3Gb/s tolerances could become a failure point at the faster 6Gb/s signal rates.

- SATA-IO recommends that only high quality cables and connectors be utilized for SATA 6Gb/s.

- Cable manufacturers and all suppliers of SATA products are encouraged to register their products on the SATA-IO Integrators list.

- This list is available to the public as a resource for determining reliable products that meet the SATA interface specification.
Enhanced Power Management

- SATA 2.6 introduced two power modes – Partial and Slumber
  - Devices can power off SATA interface circuitry when the interface is not in use, such as when a hard drive is in seek mode
- For SATA 2.6 devices, the shift from Partial to Slumber mode required the host to first change modes from Partial to Active before entering Slumber mode
- SATA 2.6 devices have to wait for permission from the host to remove power supplied to the interface
- SATA Revision 3.0 introduces Automatic Partial to Slumber mode transition
  - Eliminates the need to first enter Active mode
  - Enables either hosts or devices to initiate Slumber without asking permission from the other side
- These two improvements increase the length of time SATA interfaces can be powered down
- It is expected 6Gb/s device will need half of active time of 3Gb/s device to transfer same amount of data
Multimedia Support through Native Command Queuing (NCQ) Streaming

- NCQ streaming was designed to enable isochronous data transfer, such as audio and video, while also improving the performance of lower priority transfers.
- Host sends a time limit with a command. If a device cannot complete the command in the designated time, the device will abort the command and either returns an undefined response or truncated data.
- NCQ streaming provides a mechanism for optimizing data access, allowing the hard drive to decide the most efficient way the group of commands is to be executed.

- Freescale SATA3 controller does not support NCQ streaming.
Freescale SATA3 Controller Features

- First device with SATA3 support is LS1021A family, followed with LS2080A family.
- Designed to comply with SATA 3.0 specification
- Supports speeds: 1.5 Gbps, 3 Gbps, 6 Gbps
  
  Automatic speed negotiation
- Designed to comply with AHCI 1.3 specification
- Supports ATAPI devices
- Contains high-speed descriptor-based DMA controller
- Supports native command queuing (NCQ) commands
- Supports port multiplier operation
  
  Supports FIS-based switching
- Supports SATA BIST mode
- Supports asynchronous notification
AHCI 1.3

- AHCI stands for Advanced Host Controller Interface
- It functions at Application layer
- AHCI defines both hardware behavior and software programming interface.
- HBA Configuration Registers is related to PCI. Not implemented in Freescale SATA3 controller
AHCI 1.3 Register definitions

- AHCI supports 1-32 ports
- Port registers
  One set per port
- SATA Port number
  LS1021A: 1 port
  LS2080A: 2 port
  LS2080A actually has two SATA controllers, each AHCI controller supports one SATA port

### 3 HBA MEMORY REGISTERS

#### 3.1 Generic Host Control
- Offset 00h: CAP – HBA Capabilities
- Offset 04h: GHIC – Global HBA Control
- Offset 08h: IS – Interrupt Status Register
- Offset 0Ch: PI – Ports Implemented
- Offset 10h: VS – AHCI Version
- Offset 14h: CCC_CTL – Command Completion Coalescing Control
- Offset 18h: CCC_PORTS – Command Completion Coalescing Ports
- Offset 1Ch: EM_LOC – Enclosure Management Location
- Offset 20h: EM_CTL – Enclosure Management Control
- Offset 24h: CAP2 – HBA Capabilities Extended
- Offset 28h: BOHC – BIOS/OS Handoff Control and Status

#### 3.2 Vendor Specific Registers

#### 3.3 Port Registers (one set per port)
- Offset 00h: PxCLB – Port x Command List Base Address
- Offset 04h: PxCLBU – Port x Command List Base Address Upper 32-bits
- Offset 08h: PxFB – Port x FIS Base Address
- Offset 0Ch: PxFBU – Port x FIS Base Address Upper 32-bits
- Offset 10h: PxIS – Port x Interrupt Status
- Offset 14h: PxIE – Port x Interrupt Enable
- Offset 18h: PxCMD – Port x Command and Status
- Offset 20h: PxTFD – Port x Task File Data
- Offset 24h: PxSIG – Port x Signature
- Offset 28h: PxSSTS – Port x Serial ATA Status (SCR0: SStatus)
- Offset 2Ch: PxSCTL – Port x Serial ATA Control (SCR2: SControl)
- Offset 30h: PxSERR – Port x Serial ATA Error (SCR1: SError)
- Offset 34h: PxSACT – Port x Serial ATA Active (SCR3: SActive)
- Offset 38h: PxCI – Port x Command Issue
- Offset 3Ch: PxSNTF – Port x Serial ATA Notification (SCR4: SNotification)
- Offset 40h: PxFBS: Port x FIS-based Switching Control
- Offset 70h to 7Fh: PxVS – Vendor Specific
AHCI data structure

- HBA registers in CCSR space
AHCI data structure - cont
Native Command Queuing (NCQ) Support

- Reordering the multiple queued commands by the drive to reduce the number of rotations to complete them is called NCQ capability.

- Example: data at locations A, B, C, and D on a drive
  - Without NCQ, fetching order remains the same as A, B, C and D and this could take 21/2 rotations.
  - With NCQ, reordering the fetch to A, D, C and B the data can be read in 1 rotation.

- Because the location and organization of the data is best known to the drives, the optimal reordering of commands can be best done by the drives.
AHCI 1.3 Native command queuing (NCQ) support

- Supports 32 command slots
- Software needs to poll
  - $\text{PxCI}[31:0]$ Port $x$ Command Issue
  - $\text{PxSACT}[31:0]$ Port $x$ SATA Active

An empty command slot has its respective bit cleared to ‘0’ in both the $\text{PxCI}$ and $\text{PxSACT}$ registers.

To issue an NCQ command, the host should:

- Select an unused command slot
- Build the command table and command header, ensuring that the NCQ tag matches the slot number
- Set $\text{PxFBS.DEV}$ to the value of the Port Multiplier port field in the command header
- Set the bit in $\text{PxSACT}$ that corresponds to the command slot being used
- Set the bit in $\text{PxCI}$ that corresponds to the command slot being used
Port Multiplier

- Port multipliers are easy way of expanding data storage capacity. It attaches up to 15 devices to a host.
- Variety of port multipliers are available in the market
- Simple RAID systems could be built with single SATA host link
Freescale Supports Port Multiplier with FIS-based switching

- Port multiplier has two modes
  - Command based switching (required)
  - FIS-based switching (optional in AHCI spec)

Freescale SATA controller supports FIS-based switching which has better performance.

- To enable use of FIS-based switching, software shall set the PxFBS.EN bit to ‘1’ while PxCMD.ST is cleared to ‘0’.
SATA BIST (Build in Self Test) Mode Support

- Regular BIST support
  - Use a SATA analyzer to send BIST activate command
  - The command causes the host to enter test mode
  - The host can send out the test pattern or do loopback
  - Disconnect SATA analyzer, connect desired equipment for measurement.

## BIST Activate FIS

<table>
<thead>
<tr>
<th></th>
<th>Byte 3</th>
<th>Byte 2</th>
<th>Byte 1</th>
<th>Byte 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved</td>
<td>[TASLFPRV]</td>
<td>Reserved</td>
<td>FIS Type 58h</td>
</tr>
</tbody>
</table>

- **T**: Far end transmit only – transmit Dwords defined in words 1 & 2
- **A**: No ALIGN transmission (valid only with T)
- **S**: Bypass scrambling (valid only with T)
- **L**: Far end retimed loopback with ALIGN insertion
- **F**: Far end analog loopback
- **P**: Transmit primitives defined in words 1 & 2 of the FIS
- **R**: Reserved
- **V**: Vendor Unique Test Mode – other bits undefined
SATA BIST (Build in Self Test) Mode Support

- **Vendor Specific BIST support**
  - PHY loop back mode
    - This mode allows Tx data be looped back to Rx path
  - Use PPCFG (Port Phy1Cfg Register)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>1’h0</td>
<td>R/W</td>
<td><strong>PhyControl BIST Pattern no Aligns (PBPNA):</strong> Setting this bit will cause the Phy Ctrl Pattern generator to transmit each pattern continuously</td>
</tr>
<tr>
<td>25</td>
<td>1’h0</td>
<td>R/W</td>
<td><strong>PhyControl BIST Clear Error (PBCE):</strong> Setting this bit to 1 clears the pattern match error bit. When a pattern mismatch occurs this bit needs to be set then negated to clear the error</td>
</tr>
<tr>
<td>24</td>
<td>1’h0</td>
<td>R/W</td>
<td><strong>PhyControl BIST Pattern Enable (PBPE):</strong> Setting this bit to 1 enables the Phy Control Test Pattern generation</td>
</tr>
</tbody>
</table>
| 23:21| 3’h0 | R/W| **PhyControl BIST Pattern Select (PBPS):**
|      |      |   | 000        | LBP (Generator Only) |
|      |      |   | 001        | LFTP               |
|      |      |   | 010        | MFTP               |
|      |      |   | 011        | HFTP               |
|      |      |   | 100        | PRBS Pattern       |
|      |      |   | 101 to 111 | BIST Pattern (Default Pattern) |

Those data pattern are used in SATA compliance test.
Asynchronous Notification

• Asynchronous notification is a mechanism for a device to send a notification to the host that the device requires attention.

• Application
  – Media insertion in an ATAPI device (such as DVD ROM)
  – A hot plug event on a Port Multiplier port.

• Mechanism: To indicate that the device needs attention
  – The device shall issue a Set Devices Bits FIS to the host
    with Interrupt ‘I’ = 1
    Notification ‘N’ = 1
    Error, Status fields shall accurately reflect the current status of the device.
Power Management

SATA interface PM has three states

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
</table>
| PHYRDY  | Normal Operation mode  
Phy ready, full power, TX/RX on |
| Partial | PHY is in reduced power state  
Tx/Rx in common mode voltage  
Exit latency < 10 us |
| Slumber | PHY is in reduced power state  
Tx/Rx is allowed to float  
Exit latency < 10 ms |

SATA Protocol uses four Primitives & OOB (Out of Band) signaling for power state transition
- PMREQ_P  
Request to enter Partial mode
- PMREQ_S  
Request to enter Slumber mode
- PMACK_P  
Power management acknowledge
- PMNACK_P  
Power management not acknowledge
SATA Primitives

- A **primitive** is a single Dword (4 bytes) of information that consists of a control character in byte 0 followed by three additional data characters in bytes 1-3.

**Table 63 – Primitive Encoding**

<table>
<thead>
<tr>
<th>Primitive Name</th>
<th>Byte 3 Contents</th>
<th>Byte 2 Contents</th>
<th>Byte 1 Contents</th>
<th>Byte 0 Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALIGNₚ</td>
<td>D27.3</td>
<td>D10.2</td>
<td>D10.2</td>
<td>K28.5</td>
</tr>
<tr>
<td>CONTₚ</td>
<td>D25.4</td>
<td>D25.4</td>
<td>D10.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>DMATₚ</td>
<td>D22.1</td>
<td>D22.1</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>EOFₚ</td>
<td>D21.6</td>
<td>D21.6</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>HOLDₚ</td>
<td>D21.6</td>
<td>D21.6</td>
<td>D10.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>HOLDAₚ</td>
<td>D21.4</td>
<td>D21.4</td>
<td>D10.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>PMACKₚ</td>
<td>D21.4</td>
<td>D21.4</td>
<td>D21.4</td>
<td>K28.3</td>
</tr>
<tr>
<td>PMNAKₚ</td>
<td>D21.7</td>
<td>D21.7</td>
<td>D21.4</td>
<td>K28.3</td>
</tr>
<tr>
<td>PMREQₚ</td>
<td>D23.0</td>
<td>D23.0</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>PMREQ_Sₚ</td>
<td>D21.3</td>
<td>D21.3</td>
<td>D21.4</td>
<td>K28.3</td>
</tr>
<tr>
<td>R_ERRₚ</td>
<td>D22.2</td>
<td>D22.2</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>R_IPₚ</td>
<td>D21.2</td>
<td>D21.2</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>R_OKₚ</td>
<td>D21.1</td>
<td>D21.1</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>R_RDYₚ</td>
<td>D10.2</td>
<td>D10.2</td>
<td>D21.4</td>
<td>K28.3</td>
</tr>
<tr>
<td>SOFₚ</td>
<td>D23.1</td>
<td>D21.3</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>SYNCₚ</td>
<td>D21.5</td>
<td>D21.5</td>
<td>D21.4</td>
<td>K28.3</td>
</tr>
<tr>
<td>WTRMₚ</td>
<td>D24.2</td>
<td>D24.2</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
<tr>
<td>X_RDYₚ</td>
<td>D23.2</td>
<td>D23.2</td>
<td>D21.5</td>
<td>K28.3</td>
</tr>
</tbody>
</table>
Out-of-Band Signal Forms

Figure 160 – OOB Signals

Table 47 – OOB Signal Times

<table>
<thead>
<tr>
<th>Time</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>160 UI\textsubscript{ooob} (106.7 ns nominal)</td>
</tr>
<tr>
<td>T2</td>
<td>480 UI\textsubscript{ooob} (320 ns nominal)</td>
</tr>
</tbody>
</table>
Out-of-Band Signal Usage Example

Figure 167 – Power-On Sequence
Power Management: Power State Transition
For SATA 1.5Gbps and 3.0Gbps

- Entering PARTIAL or SLUMBER
  - Host or device may send PMREQ_P or PMREQ_S
- To resume from PARTIAL or SLUMBER, either host or device
  - Sends COMWAKE OOB, Receives COMWAKE response, and moves to Phy ready.
  - Resumes right where it left off
- No path between PARTIAL and SLUMBER
SATA 6.0 Gbps: Automatic Partial to Slumber Transition (APST)

- Host and device may be independently aware of conditions where they may transition to the Slumber state to save power without impacting performance (e.g., idle, seeking).

- APST: PHY transits to Slumber from Partial without first entering PHYRDY.

- By default, the feature is disabled on power-up. SET FEATURES command: enable/disable the feature.

- If enable, once host and device enter PARTIAL mode, then either host or device can enter SLUMBER independently.

- Host has no knowledge of the state of device (PARTIAL or SLUMBER) and vice versa. One side has to tolerate the max latency of slumber exit time for the other side.
SATA 6.0 Gbps: Automatic Partial to Slumber Transition (APST)

- **IDENTIFY DEVICE** command, device communicate its info to host
- Word 76, bit 14, 13 are relate to Automatic Partial to Slumber Transitions
- Host Automatic Partial to Slumber shall only be used if the device support host Automatic Partial to Slumber transition (Word 76, bit 13)
- Device Automatic Partial to Slumber shall only be used if
  - Host can support
  - Enabled device Automatic Partial to Slumber transitions (via Set Features)
- Slumber transitions (via Set Features)

<table>
<thead>
<tr>
<th>Word 76</th>
<th>Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>15 Supports READ LOG DMA EXT as equivalent to READ LOG EXT</td>
</tr>
<tr>
<td>F</td>
<td>14 <strong>Supports Device Automatic Partial to Slumber transitions</strong></td>
</tr>
<tr>
<td>F</td>
<td>13 <strong>Supports Host Automatic Partial to Slumber transitions</strong></td>
</tr>
<tr>
<td>F</td>
<td>12 Supports Native Command Queuing priority information</td>
</tr>
<tr>
<td>F</td>
<td>11 Supports Unload while NCQ commands outstanding</td>
</tr>
<tr>
<td>F</td>
<td>10 Supports Phy event counters</td>
</tr>
<tr>
<td>F</td>
<td>9 Supports receipt of host-initiated interface power management requests</td>
</tr>
<tr>
<td>F</td>
<td>8 Supports Native Command Queuing</td>
</tr>
<tr>
<td>F</td>
<td>7-4 Reserved for future Serial ATA signaling speed grades</td>
</tr>
<tr>
<td>R</td>
<td>3 1 = Supports Serial ATA Gen3 signaling speed (6.0 Gbps)</td>
</tr>
<tr>
<td>F</td>
<td>2 1 = Supports Serial ATA Gen2 signaling speed (3.0 Gbps)</td>
</tr>
<tr>
<td>F</td>
<td>1 1 = Supports Serial ATA Gen1 signaling speed (1.5 Gbps)</td>
</tr>
<tr>
<td>F</td>
<td>0 Shall be cleared to zero</td>
</tr>
</tbody>
</table>
Power Management: AHCI 1.3

3.3.7 Offset 18h: PxCMD – Port x Command and Status

<table>
<thead>
<tr>
<th>Offset (Hex)</th>
<th>RW/RO</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>RW/RO</td>
<td>0</td>
<td>Aggressive Link Power Management Enable (ALPE): When set to ‘1’, the HBA shall aggressively enter a lower link power state (Partial or Slumber) based upon the setting of the ASP bit. Software shall only set this bit to ‘1’ if CAP.SALP is set to ‘1’; if CAP.SALP is cleared to ‘0’ software shall treat this bit as reserved. See section 8.3.1.3 for details.</td>
</tr>
<tr>
<td>27</td>
<td>RW/RO</td>
<td>0</td>
<td>Aggressive Slumber / Partial (ASP): When set to ‘1’, and ALPE is set, the HBA shall aggressively enter the Slumber state when it clears the PxCI register and the PxsSACT register is cleared or when it clears the PxsSACT register and PxCi is cleared. When cleared, and ALPE is set, the HBA shall aggressively enter the Partial state when it clears the PxCi register and the PxsSACT register is cleared or when it clears the PxsSACT register and PxCi is cleared. If CAP.SALP is cleared to ‘0’ software shall treat this bit as reserved. See section 8.3.1.3 for details.</td>
</tr>
</tbody>
</table>

- ALPE=1, ASP=1
  SATA host controller shall enter slumber mode if command slots are empty (both PxCi and PxsSACT=0)

- ALPE=1, ASP=0
  SATA host controller shall enter partial mode if command slots are empty (both PxCi and PxsSACT=0)
Introducing The QorIQ LS2 Family

Breakthrough, software-defined approach to advance the world’s new virtualized networks

New, high-performance architecture built with ease-of-use in mind
Groundbreaking, flexible architecture that abstracts hardware complexity and enables customers to focus their resources on innovation at the application level

Optimized for software-defined networking applications
Balanced integration of CPU performance with network I/O and C-programmable datapath acceleration that is right-sized (power/performance/cost) to deliver advanced SoC technology for the SDN era

Extending the industry’s broadest portfolio of 64-bit multicore SoCs
Built on the ARM® Cortex®-A57 architecture with integrated L2 switch enabling interconnect and peripherals to provide a complete system-on-chip solution
QorIQ LS2 Family

Key Features

Unprecedented performance and ease of use for smarter, more capable networks

High performance cores with leading interconnect and memory bandwidth
- 8x ARM Cortex-A57 cores, 2.0GHz, 4MB L2 cache, w Neon SIMD
- 1MB L3 platform cache w/ECC
- 2x 64b DDR4 up to 2.4GT/s

A high performance datapath designed with software developers in mind
- New datapath hardware and abstracted acceleration that is called via standard Linux objects
- 40 Gbps Packet processing performance with 20Gbps acceleration (crypto, Pattern Match/RegEx, Data Compression)
- Management complex provides all init/setup/teardown tasks

Leading network I/O integration
- 8x1/10GbE + 8x1G, MACSec on up to 4x 1/10GbE
- Integrated L2 switching capability for cost savings
- 4 PCIe Gen3 controllers, 1 with SR-IOV support
- 2 x SATA 3.0, 2 x USB 3.0 with PHY

SDN/NFV Switching

Data Center

Wireless Access
See the LS2 Family First in the Tech Lab!

4 new demos built on QorIQ LS2 processors:

- Performance Analysis Made Easy
- Leave the Packet Processing To Us
- Combining Ease of Use with Performance
- Tools for Every Step of Your Design