LIBRARY
LIBRARY OF CONGRESS

Office of Business Enterprises
Duplication Services Section

THIS IS TO CERTIFY that the collections of the Library of Congress contain a book entitled ***High Speed Serdes Devices and Applications*** (New York: Springer, 2008) by David R. Stauffer, Jeanne Trinko Mechler, Michael Sorna, Kent Dramstad, Clarence R. Ogilvie, Amanullah Mohammad, and James Rockrohr and that the attached photocopies—the spine, front cover, first title page, second title page spread, publication information page, table of contents, pages 1 through 19, 41 through 52, pages 58 through 59, 116 through 118, 165 through 168, 197 through 201, 213 through 216, and back cover—are a true representation from that work.

THIS IS TO CERTIFY FURTHER, that the spine bears a Library of Congress call number that reads "TK 7870 .H52 2008 Copy 1" and a label that reads "CBN BRANCH GenColl, that the publication information page bears a handwritten notation reading "TK 7870 .H52 2008 Copy 1," and that the back cover bears a label that reads "LC Copyright 0 030 716 497 8."

IN WITNESS WHEREOF, the seal of the Library of Congress is affixed hereto on June 4, 2025.

Rosalina Delgado-Jones
Assistant Business Enterprises Officer
Office of Business Enterprises
Library of Congress




101 Independence Avenue, SE Washington, DC 20540–4917 Tel 202.707.5650 www.loc.gov;
duplicationservices@loc.gov

Marvell Semiconductor, Inc. - Ex. 1014, Page 000002

# High Speed Serdes Devices and Applications

Marvell Semiconductor, Inc. - Ex. 1014, Page 000003

David R. Stauffer • Jeanne Trinko Mechler
Michael Sorna • Kent Dramstad
Clarence R. Ogilvie • Amanullah Mohammad
James Rockrohr

# High Speed Serdes Devices and Applications




Marvell Semiconductor, Inc. - Ex. 1014, Page 000005

TK7870
.H52
2008
Cage 1

David R. Stauffer
IBM Corporation
Essex Junction, VT
USA

Kent Dramstad
IBM Corporation
Essex Junction, VT
USA

Amanullah Mohammad
IBM Corporation
Research Triangle Park, NC
USA

Michael A. Sorna
IBM Microelectronics
Hopewell Junction, NY
USA

Jeanne T. Mechler
IBM Corporation
Essex Junction, VT
USA

Clarence R. Ogilvie
IBM Corporation
Essex Junction, VT
USA

James D. Rockrohr
IBM Microelectronics
Hopewell Junction, NY
USA

ISBN 978-0-387-79833-2 e-ISBN 978-0-387-79834-9

Library of Congress Control Number: 2008925643

© 2008 Springer Science+Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or heareafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.
While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Marvell Semiconductor, Inc. - Ex. 1014, Page 000006

# Table of Contents


Preface v
Acknowledgments vii

**Chapter 1: Serdes Concepts . . . . . 1**
1.1 The Parallel Data Bus 1
1.2 Source Synchronous Interfaces 2
Reducing the Number of I/O Pins 2
Clock Forwarding 3
Higher Speed Source Synchronous Interfaces 4
1.3 High-Speed Serdes 8
Serializer / Deserializer Blocks 9
Equalizers 10
Clock and Data Recovery (CDR) 14
Differential Driver 15
Differential Receiver 17
Diagnostic Functions 17
Phase-Locked Loop 19
1.4 Signal Integrity 19
The Channel 19
Package Models 21
Jitter 21
Channel Analysis Tools 23
1.5 Signaling Methods 24
1.6 Exercises 27

**Chapter 2: HSS Features and Functions . . . . . 31**
2.1 HSS Core Example: HSS EX10 10-Gbps Core 31
HSS EX10 Input/Output Pin Descriptions 32
HSS EX10 Register Descriptions 41
2.2 HSS EX10 Transmitter Slice Functions 53
Transmitter Parallel Data 54
Transmitter Signal Characteristics 56
Transmitter FFE Programming 58
Transmitter Power Control 59
Half-Rate/Quarter-Rate/Eighth-Rate Operation 60
JTAG 1149.1 and Bypass Mode Operation 62
PRBS / Loopback Diagnostic Features 64
Out of Band Signalling Mode (OBS) 65
Features to Support PCI Express 65
2.3 HSS EX10 Receiver Slice Functions 66
Receiver Data Interface 68
DFE and Non-DFE Receiver Modes 70




Marvell Semiconductor, Inc. - Ex. 1014, Page 000008

Serial Data Termination and AC/DC Coupling 71
Signal Detect 71
Receiver Power Control 72
JTAG 1149.1/1149.6 and Bypass Mode Operation 73
Half-Rate/Quarter-Rate/Eight-Rate Operation 76
PRBS / Loopback Diagnostic Features 77
Phase Rotator Control/Observation 78
Support for Spread Spectrum Clocking 78
Eye Quality 79
SONET Clock Output 80
Features to Support PCI Express 80
2.4 Phase-Locked Loop (PLL) Slice 80
Reference Clock 81
Clock Dividers 82
Power On Reset 82
VCO Coarse Calibration 83
PLL Lock Detection 83
Reset Sequencer 84
HSS Resynchronization 84
PCI Express Power States 87
2.5 Reset and Reconfiguration Sequences 87
Reset and Configuration 87
Changing the Transmitter Configuration 90
Changing the Receiver Configuration 92
2.6 References and Additional Reading 93
2.7 Exercises 94

**Chapter 3: HSS Architecture and Design. . . . . 99**
3.1 Phase Locked Loop (PLL) Slice 100
PLL Macro 101
Clock Distribution Macro 102
Reference Circuits 103
PLL Logic Overview 105
3.2 Transmitter Slice 107
Feed Forward Equalizer (FFE) Operation 109
Serializer Operation 112
3.3 Receiver Slice 114
Clock and Data Recovery (CDR) Operation 116
Decision Feedback Equalizer (DFE) Architectures 118
Data Alignment and Deserialization 121
3.4 References and Additional Reading 122
3.5 Exercises 123

Marvell Semiconductor, Inc. - Ex. 1014, Page 000009

**Chapter 4: Protocol Logic and Specifications ..... 125**
4.1 Protocol Specifications 125
Protocol Layers 125
Serial Data Specifications 126
Basic Concepts 132
4.2 Protocol Logic Functions 134
Bit/Byte Order and Striping/Interleaving 134
Data Encoding and Scrambling 136
Error Detection and Correction 143
Parallel Data Interface 147
Bit Alignment 152
Deskewing Multiple Serial Data Links 153
4.3 References and Additional Reading 158
4.4 Exercises 159

**Chapter 5: Overview of Protocol Standards ..... 165**
5.1 SONET/SDH Networks 168
System Reference Model 169
STS-1 Frame Format 170
STS-N Frame Format 174
Clock Distribution and Stratum Clocks 176
5.2 OIF Protocols 177
System Reference Model 177
SFI-5.2 Implementation Agreement 180
SPI-S Implementation Agreement 184
CEI-P Implementation Agreement 188
Electrical Layer Implementation Agreements 190
5.3 Ethernet Protocols 197
Physical Layer Reference Model 198
Media Access Control (MAC) Layer 201
XGMII Extender Sublayer (XGXS) 204
10-Gb Serial Electrical Interface (XFI) 207
Backplane Ethernet 213
PMD Sublayers for Electrical Variants 218
5.4 Fibre Channel (FC) Storage Area Networks 220
Storage Area Networks (SANs) 220
Fibre Channel Protocol Layers 222
Framing and Signaling 222
Physical Interfaces 229
10-Gbps Fibre Channel 236
5.5 PCI Express 237
PCI Express Architecture 238
Physical Layer Logic 241
Electrical Physical Layer 246
Power States 249
PCI Express Implementation Example 250

Marvell Semiconductor, Inc. - Ex. 1014, Page 000010

5.6 References and Additional Reading 251
5.7 Exercises 254

**Chapter 6: Reference Clocks . . . . . 263**
6.1 Clock Distribution Network 263
Single-Ended vs. Differential Reference Clocks 263
Reference Clock Sources 265
Special Timing Requirements 268
Special Test Requirements 270
6.2 Clock Jitter 270
Jitter Definitions 271
Jitter Effects 276
PLL Jitter 277
6.3 Clock Floorplanning 281
Clock Tree Architecture 281
Clock Tree Wiring 282
6.4 Signal Integrity of the Clock Network 283
Analog Signal Levels and Slew Rates 283
Duty Cycle Distortion 286
Differential Clock Analysis Methodology 288
6.5 References and Additional Reading 293
6.6 Exercises 293

**Chapter 7: Test and Diagnostics . . . . . 297**
7.1 IEEE JTAG 1149.1 and 1149.6 298
JTAG 1149.1 Overview 299
HSS Core Support for JTAG 1149.1 302
HSS Core Support for JTAG 1149.6 303
7.2 PRBS Testing and Loopback Paths 306
Loopback Paths 306
PRBS Circuits and Data Patterns 309
PRBS Test Sequence 314
7.3 Logic Built-In-Self-Test (LBIST) 317
LBIST Architecture 317
LBIST Considerations for HSS Cores 319
7.4 Manufacturing Test 320
Chip Level Test 320
HSS Macro Test 324
7.5 Characterization and Qualification Testing 327
Transmitter Tests 328
Receiver Tests 335
General Tests 338
7.6 References and Additional Reading 340
7.7 Exercises 340

Marvell Semiconductor, Inc. - Ex. 1014, Page 000011

**Chapter 8: Signal Integrity ..... 345**
8.1 Probability Density Functions 345
Gaussian Distribution 345
Dual-Dirac Distribution 348
8.2 Jitter 349
Jitter Components 349
Deterministic Jitter 352
Random Jitter 356
Total Jitter and Mathematical Models 358
Jitter Budgets 362
Jitter Tolerance 364
8.3 Spice Models 365
Traditional Spice Models 365
Hybrid Spice/Behavioral Models 367
Spice Simulation Matrices 369
8.4 Statistical Approach to Signal Integrity 372
Analysis Approach 373
HSSCDR Software 388
8.5 References and Additional Reading 393
8.6 Exercises 394

**Chapter 9: Power Analysis. ..... 397**
9.1 Digital Logic Circuits 397
Digital Logic Active or AC Power 397
Digital Logic Leakage or DC Power 402
9.2 Non Digital Logic Circuits 410
AC (Active) Power 410
DC (Leakage) Power 410
Quiescent Power 410
9.3 HSS Power 411
HSS Power Equation 411
Multiple Power Supplies 412
Chip Fabrication Process 413
Mode-Dependent Power 414
Power Dissipation Breakdown 416
9.4 Reducing Power Dissipation 417
Power Concerns for the HSS Core Design 417
Power Dissipation Concerns for the Chip Designer 420
9.5 References and Additional Reading 421
9.6 Exercises 421

Marvell Semiconductor, Inc. - Ex. 1014, Page 000012

**Chapter 10: Chip Integration ..... 425**
10.1 Simulation Models 427
Reset and Initialization Short Cuts 427
Simulation 'X' States 429
Modeled and Unmodeled Behavior 432
10.2 Test Synthesis 434
Scan Test Support 435
Macro Test Support 436
JTAG Logic Connections 440
Automation of Test Requirements 442
Running Macro Test using the JTAG Interface 444
10.3 Static Timing Analysis 445
Clock Timing 445
Receiver Parallel Data Outputs 450
Register Interface 452
Transmitter Synchronization 454
Serial Data Timing 456
Skew Management 457
Timing Backannotation for Simulation 458
10.4 Chip Floorplan and Package Considerations 459
Packages 459
Chip Physical Design 466
10.5 References 471
10.6 Exercises 472

**Index..... 475**

Marvell Semiconductor, Inc. - Ex. 1014, Page 000013

# Chapter 1
# Serdes Concepts

This chapter describes basic methods of transferring data from one chip to another chip, either on the same circuit board or across a cable or backplane to another circuit board. After reading this chapter, the reader should have a basic understanding of the rationale for using high-speed serializer/deserializer (Serdes) devices, and the inherent problems introduced by the high-speed operation of such devices.

## 1.1 The Parallel Data Bus

The simplest method of transferring data through the inputs or outputs of a silicon chip is to directly connect the datapath from one chip to the next chip (see Fig. 1.1). Since data often consists of more than one bit of information, the datapath is more than one-bit wide. In the figure, an *n*-bit datapath inside *Chip #1* is driven through chip outputs, across an *n* bit interconnect, through inputs of *Chip #2*, to an *n*-bit datapath inside the receiving chip. Synchronous data is transferred between the two chips since both chips are clocked by the same clock source.

There are two inherent problems of the parallel data bus shown in Fig. 1.1. The first problem is that *n* input/output (I/O) pins are required on each chip to transfer the data. At one point in history this was acceptable. However, Moore's Law has driven substantial increases in the number of circuits that can be manufactured on a chip compared to a few decades ago. The pin densities of chip packaging technologies have not increased at the same pace as silicon density. Therefore, I/O pins are substantially more expensive than silicon circuits, and dedicating *n* I/O pins for the above data bus is not acceptable for most chip applications.

The second inherent problem involves meeting timing requirements. The data is launched synchronously by *Chip #1* and is captured synchronously in *Chip #2* using the same clock. The data at the inputs of *Chip #2* must meet setup and hold times relative to the clock input of the chip. These setup and hold times must be calculated with sufficient margin to allow for differences in delay of the clock distribution path to the two chips, and through the chips to the launch and capture flip-flops. Delays may vary based on chip process, voltage, and temperature (PVT) conditions, and margin must be added to account for worst case variations. For higher clock frequencies, it may be necessary to use phase-locked loops (PLLs) in the chips to adjust the clock phase in order to compensate for the clock distribution delay within the chip and adapt to changing process, voltage, and temperature conditions. If the clock frequency is high enough, it will not be possible to build a system that will reliably transfer the data across this data bus.

D. R. Stauffer et al., *High Speed Serdes Devices and Applications*.
© Springer 2008


Marvell Semiconductor, Inc. - Ex. 1014, Page 000015

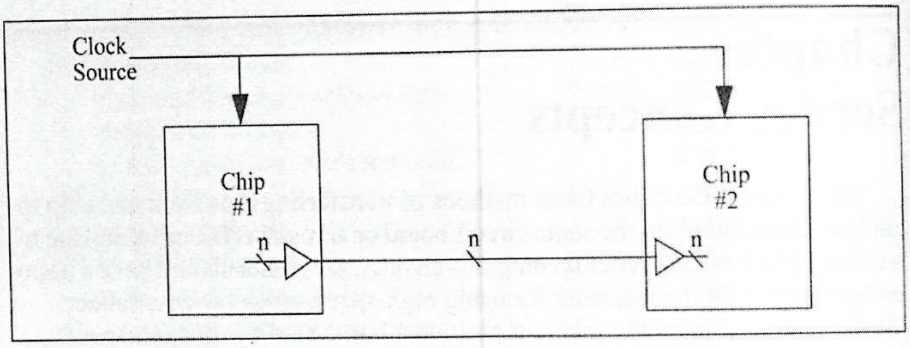



**Fig. 1.1** Parallel data bus between two chips

## 1.2 Source Synchronous Interfaces

The two problems with the parallel data bus in Sect. 1.1 can be eliminated with the modifications to the system which are discussed in this section. These approaches are extensions of the parallel data bus. The parallel data bus and all of the extensions described in this section are considered to be *source synchronous interface* architectures. Such architectures include any interface where a clock input exists that can be used to capture the received data. This may be either a reference clock used by both the transmitting and receiving chip or the transmitting chip drives a clock to the receiving chip. In either case, clock recovery circuits are *not* required for source synchronous interfaces.

### *1.2.1 Reducing the Number of I/O Pins*

The first issue to be addressed is reducing the number of I/O pins required to transfer the data between the chips. This is accomplished by multiplexing the $n$ bits of data at the output of *Chip #1* onto $k$ bits of interconnect ( $k < n$ ), and then demultiplexing the $k$ bits of interconnect at the input of *Chip #2* onto an $n$ bit internal datapath. This is shown in Fig. 1.2. The resulting system only requires $k$ I/O pins on each chip rather than the $n$ pins previously required.

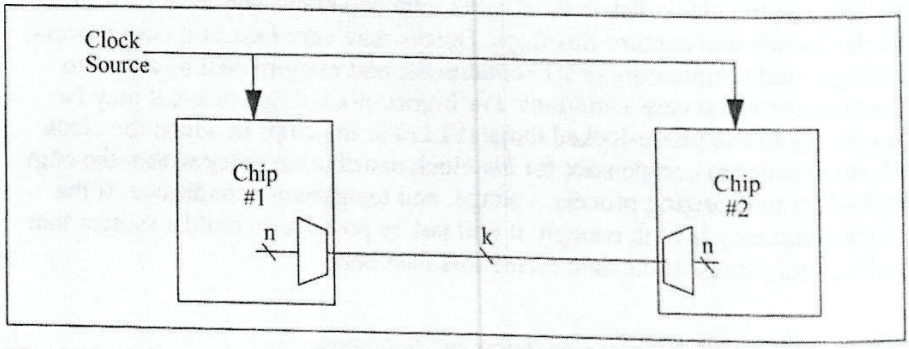



**Fig. 1.2** Serializing the data to reduce pin counts

Marvell Semiconductor, Inc. - Ex. 1014, Page 000017

Of course, while the pin count requirements have been reduced by the ratio of $k : n$, the required frequency of the reference clock has increased by the inverse of this ratio. System designers generally do not like to distribute high-speed reference clocks within the system due to noise, electromagnetic interference (EMI), and power dissipation concerns. Often, a lower frequency clock is distributed, and PLLs in the chips are used to multiply this reference clock to a usable frequency. Variability of the phase of the resulting clock, along with the higher frequency of data transfer, tends to exacerbate the timing issues of the parallel data bus approach.

## *1.2.2 Clock Forwarding*

In Fig. 1.3, a high-speed clock has been added to the datapath between the two chips. This clock source is assumed to supply a clock frequency somewhat lower than the frequency required to clock the data flip-flops on the chip interconnect. PLLs are used in each chip to generate clocks at a multiple of this frequency. The resulting clocks are used to launch and capture data in the respective chips. The output clock of the PLL in *Chip #1*, which is used to launch the data from this chip, is also an output of this chip. This clock is used by *Chip #2* to capture the data. This approach is called *clock forwarding*.

The advantage of this approach is that the high-speed clock used to launch the data at *Chip #1* is available to *Chip #2* as a reference to capture the data. Any variations in delays through clock distribution network driving the two chips does not need to be taken into account in timing analysis. Only delay variations between the clock path and the data bits are relevant. Variations between these paths due to process, voltage, and temperature track each other to some extent. The result is that timing analysis of the interface requires less margin and setup and hold times are therefore easier to meet.

So far we have not made any distinction or recommendations regarding the frequency of the high-speed clock relative to the bit rate of the interface. In general, the high-speed clock shown in the figure could be *single data rate* (SDR) or *double data rate* (DDR) (Fig. 1.4). The receiving chip captures data on every rising (or every falling) edge of an SDR clock; while the receiving chip captures data on every edge (both rising and falling edges) of a DDR clock.

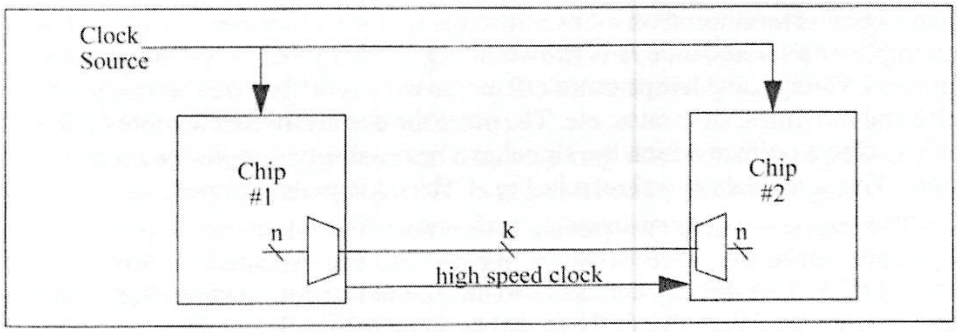



**Fig. 1.3** High-speed clock forwarded with the data

Marvell Semiconductor, Inc. - Ex. 1014, Page 000018

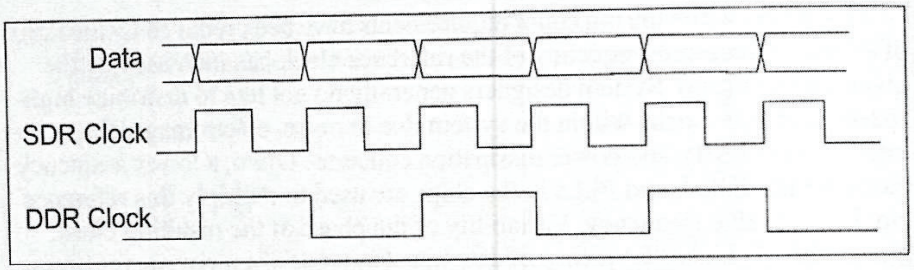



**Fig. 1.4** Single data rate and double data rate clocks

The advantage of DDR clocks over SDR clocks lies in the bandwidth requirements for the corresponding I/O drivers and receivers. An I/O cell being used at a bit rate of *b* Mbits per second requires a bandwidth sufficient to transmit a 101010... data pattern. This corresponds to a frequency spectrum with an upper fundamental frequency limit of *b*/2 MHz. The corresponding frequency of an SDR clock is *b* MHz, twice the spectral limit of the data. However, the frequency of a DDR clock for the same interface is only *b*/2 MHz, consistent with the frequency spectrum of the data. Therefore, the same I/O drivers and receivers can be used to drive and receive both the data and the DDR clock.

Regardless of whether the high-speed clock is a SDR or a DDR clock, the receiving chip uses this clock to directly capture the data. This chip also uses the reference clock to generate an internal system clock at the same frequency. These clocks are mesochronous. While the frequency is the same (given that they share a common frequency reference), the phase relationship between the clocks is unknown and may vary due to PVT variations. Therefore, the receiving chip usually retimes the received data from the interface clock domain to the clock domain of the internal chip clock. FIFOs are used to perform this retiming function. It is desirable to minimize the number of flip-flops being clocked by the interface clock in order to minimize delay in the clock distribution network; otherwise timing issues will be exacerbated.

### *1.2.3 Higher Speed Source Synchronous Interfaces*

The window of time during which data bits can be assumed to be valid is called the *eye*. This name originates from the shape of the waveform when the data signal is monitored on an oscilloscope that is continuously triggered. An example of a serial data eye is shown in Fig. 1.5a. *Eye closure* results from process, voltage, and temperature effects, as well as differences between signal rise and fall times, slew rates, etc. The more the eye is closed, the more difficult it is to find a point at which the signal can be reliability sampled to receive the data. The serial data eye shown in Fig. 1.5b is completely closed.

The largest possible *eye opening* is desirable. The width and height of an open eye can be measured as shown in Fig. 1.5c. The expected bit error rate (BER) of the link directly correlates to the amount of eye opening (both width and height). This section briefly describes some approaches to minimize eye closure of the data signal. Eye waveforms are discussed further in later chapters.

Marvell Semiconductor, Inc. - Ex. 1014, Page 000019

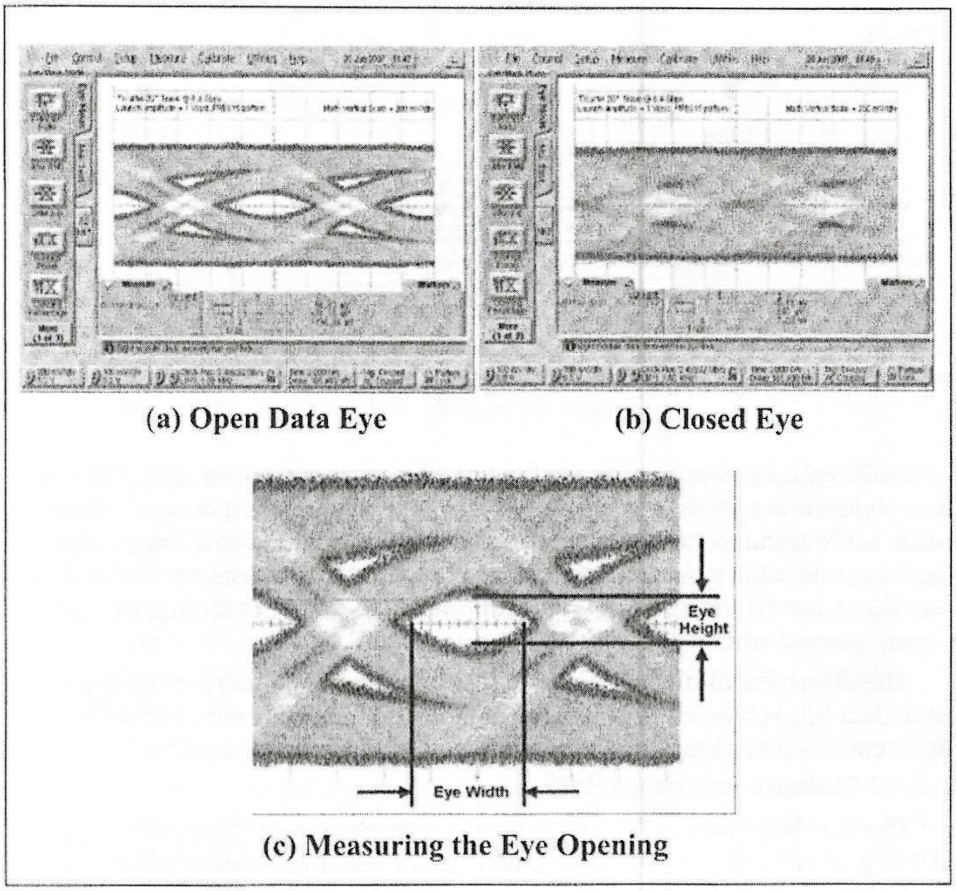



**Fig. 1.5** Example of a serial data eye

### 1.2.3.1 Differential Signals

Unequal signal rise and fall times of nondifferential signals contribute to eye closure. Signals switching on the chip also create current variations on the power distribution grid of the chip, which in turn cause variations in voltage drop (noise) that can cause variation in delays of surrounding circuits. One method of reducing the effects of these phenomenon on the eye width is to drive differential signals between chips.

Differential signals represent the data bit using two electrical signals (*true* and *complement* signals). A logic "0" is represented by the *true* signal driven to its lower voltage limit, and the *complement* signal driven to its upper voltage limit; a logic "1" is represented by the *true* signal driven to its upper voltage limit, and the *complement* signal driven to its lower voltage limit. A differential receiver device interprets the logic bit value based on the difference between the two signals, and not based on the level of either signal individually.

Marvell Semiconductor, Inc. - Ex. 1014, Page 000020

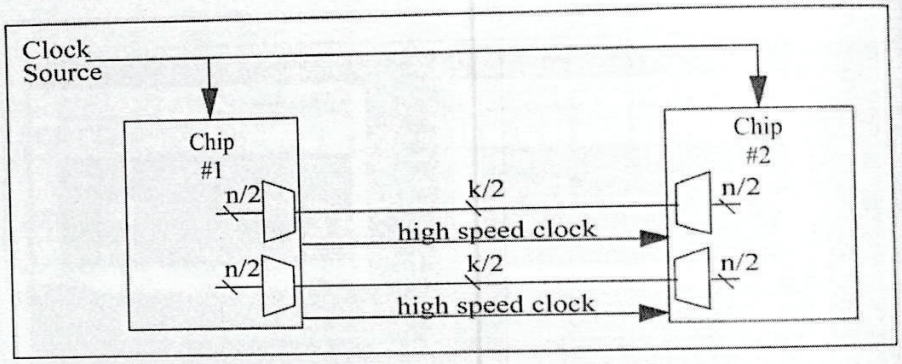



**Fig. 1.6** Multiple sets of data with separate high-speed clocks

Differential driver circuits tend to have linear current draw and generate less noise on the power supply than equivalent single-ended drivers. Most noise sources induce voltage variation equally on both the true and complement signals; such *common mode noise* is ignored by the receiver. Also, since one leg of the differential signal is rising while the other is falling, or vice versa, unequal rise and fall time effects cancel.

The drawback of differential signals is that two chip pins are required for each data bit. However, this is offset by the higher speeds possible with differential signals that are not possible with single-ended signals.

### 1.2.3.2 Multiple Interface Clocks

The interface clock in Fig. 1.3 is the same clock as is used to launch the data, and in general is driven from a point in the clock distribution network as close to the actual flip-flops that launch the data as possible. Phase variation is introduced by any circuits which are not common to both the data path and clock path. Silicon process variables do vary from circuit to circuit on the same chip, the power distribution network may have unequal voltage drops to different circuits which may vary based on switching currents, and the temperature may vary from point to point on the chip. Tolerances and limits for all of these parameters must be taken into account when calculating delays, setup times, and hold times necessary for correct capture of the received data. At higher bit/baud rates, these parameters may significantly reduce the eye opening, and become the dominant mechanism for limiting the speed of the interface.

To maximize the eye width, the path through the clock tree to each of the data flip-flops and to the clock output should share as many circuits as possible, and the output driver for the clock should be similar to the output drivers for the data. Ideally, the same clock buffer should drive the clock to the output driver and should drive the clock input to all of the data flip-flops. The larger the number of bits in the data bus, the more difficult this becomes to implement. I/O drivers must be physically distributed based on the groundrules for connections to package pins. The greater the distance between circuits, the more process, voltage, and temperature variation, and the more circuits in the clock distribution network which cannot be shared due to lack of proximity.

Marvell Semiconductor, Inc. - Ex. 1014, Page 000021

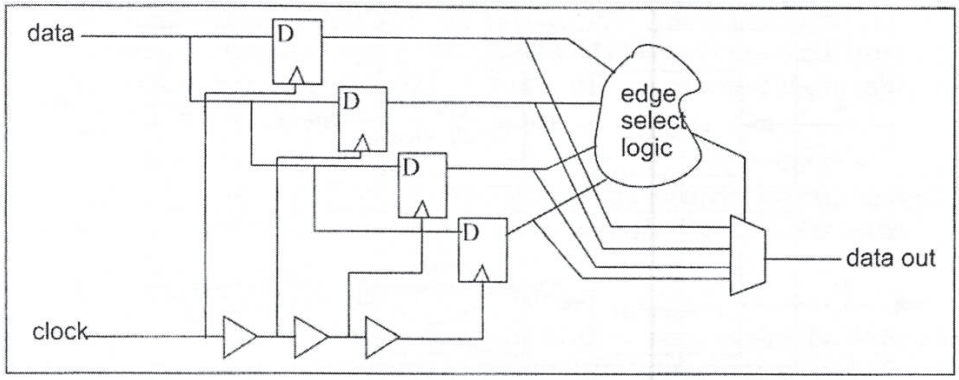



**Fig. 1.7** Adapting the sampling clock phase in the receiver

One technique used to improve eye width is to limit the number of data bits associated with a given interface clock line. Wider data busses are built by using multiple interface clocks, each clock associated with a subset of the data bits. An example of this is shown in Fig. 1.6, where the *k* bit interconnect has been subdivided into two groups, each with its own high-speed interface clock. Note that the receiving chip must capture each group of data bits in separate clock domains, and needs to retime this data to the common clock domain internal to the chip.

### 1.2.3.3 Sample Edge Adaptation

Another technique used to permit higher speed operation of source synchronous interfaces is to process the data signal at the receiver and adapt the sampling phase of the clock on a per-bit basis. This is done by connecting the received interface clock signal to the input of a multitap delay line, and capturing the data signal in multiple flip-flops clocked by different clock phases. Logic can then be used to determine the clock phases between which data transitions are occurring, and select the optimal clock phase to be used to capture the data. This scheme is shown in Fig. 1.7.

Schemes, such as shown in Fig. 1.7, may require a training pattern either upon initialization of the interface or at regular intervals. If a training pattern is used, phase selections remain static between training periods. More complex implementations adjust dynamically based on the received data or based on training patterns embedded in the data stream. Alternative architectures which apply the data to the delay line are also possible. Note, however, that an inherent characteristic of most of these schemes is that the phase adjustment is less than plus/minus one bit time, and there must be a sufficient eye opening such that an optimal sampling phase exists.

Given the advanced schemes discussed above, data rates for source synchronous interfaces can be extended to several Gigabits per second (Gbps) per interconnect bit. However, PVT variations make further increases in interface speeds prohibitively complex. Beyond these speeds, High-Speed Serdes devices that extract the clock from edge transitions in the data stream become the preferred solution.

Marvell Semiconductor, Inc. - Ex. 1014, Page 000022

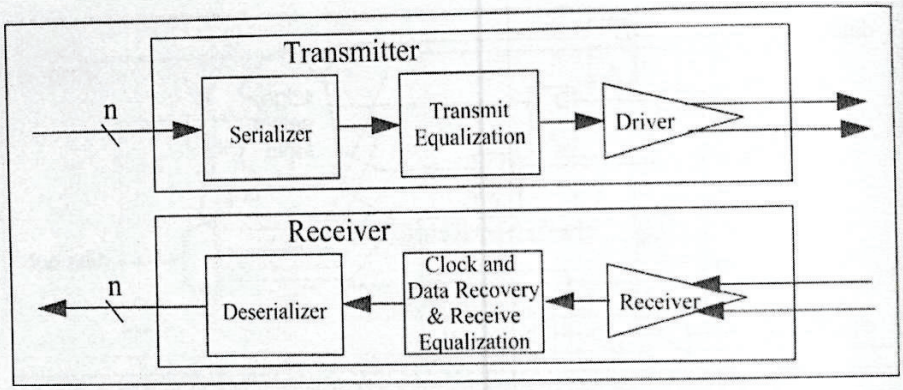



**Fig. 1.8** Basic block diagram of typical high-speed serdes

## 1.3 High-Speed Serdes

High-Speed Serializer/Deserializer (HSS) devices are the dominant implementation of I/O interfaces at speeds of 2.5 Gbps and higher. Such devices are differentiated from source–synchronous interfaces in that the receiver device contains a clock and data recovery (CDR) circuit which dynamically determines the optimal sampling point of the data signal based upon the transition edges of the signal. In other words, clock information is extracted directly from the data rather than relying on a separate clock.

Figure 1.8 illustrates the basic block diagram of the transmit and receive channels of an HSS device. The transmitter serializes parallel data, equalizes it for reasons that will be explained shortly, and then drives the serial data onto a differential signal pair of interconnect wires. Feed forward equalizers (FFE) are commonly used in High-Speed Serdes devices, as discussed in Sect. 1.3.2. The receiver consists of a differential receiver, a CDR circuit which may also integrate an equalizer, and deserializes the data based upon the sample point established by the CDR. Peaking amplifiers and/or decision feedback equalizers (DFE) are commonly used for equalization in High-Speed Serdes receiver devices.

Note that Serdes cores are often designed to group multiple transmit and/or receive channels into a single device. The individual channels generally operate independently. Grouping channels allow some circuits to be shared across channels (for example the PLL noted below), and therefore the resulting block is more efficient in terms of chip area, cost, and power.

Serdes cores which contain only transmit or only receive channels are called *simplex* cores; Serdes cores which contain both transmit and receive channels are called *full duplex* cores. Note that the terminology "full duplex" does not imply that the electrical interface is bidirectional. Any given electrical interconnect channel has a fixed direction of data transmission. If a protocol application requires "full duplex" communication, then independent transmit and receive channels with independent interconnections are used to implement the interface. Rationale for using simplex vs. full duplex cores may include

Marvell Semiconductor, Inc. - Ex. 1014, Page 000024

(1) chip floorplan to minimize wiring crossings in the package design or circuit board design; (2) signal integrity concerns due to near-end crosstalk from transmit signals onto receive signals; (3) or applications where the number of transmit and receive channels is not equal.

The remainder of this section generically describes various circuits mentioned above in more detail, as well as providing generic descriptions of other circuits and functions commonly found in High-Speed Serdes cores.

## *1.3.1 Serializer/Deserializer Blocks*

Conceptually, the input to the serializer transmit stage is an *n*-bit datapath which is serialized to a one-bit serial data signal for application to the FFE and Driver stages. Generally the value of *n* is a multiple of 8 or 10, and may be programmable on some implementations. Values of *n* which are multiples of 8 are useful for sending unencoded and/or scrambled data bytes; values of *n* which are multiples of 10 are useful for protocols which use 8B/10B coding, as discussed further in Sect. 4.2.2.1. (The 8B/10B encoder is generally implemented by logic outside the Serdes core.)

For simplicity, the block diagram in Fig. 1.8 illustrates the serializer feeding one-bit data into the transmit equalization block. Actual implementations may vary, and this datapath may be one or more bits wide. A wider datapath through the equalizer block results in a more complex design, but requires a lower operating frequency. Some implementations may initially multiplex the *n*-bit input to an *m*-bit datapath ( $m < n$ ) prior to the equalizer, and perform the remainder of the serialization at the driver stage.

The serializer stage latches data on the *n*-bit input at the frequency of *baud rate*/*n*. The high-speed clock in the Serdes is divided down to generate a sample clock for the parallel data. Because the phase of this clock is determined by the internal state of the serializer, the Serdes channel generally provides this clock as an output for use by logic driving data to the transmit channel.

Conceptually, the deserializer receive block performs the inverse function of the serializer block. Serial data is deserialized onto an *n*-bit databus of similar width to the serializer. A sample clock is generated by dividing down the internal high-speed clock, and this clock is supplied as an output for use by logic latching the parallel data. In a similar manner to the serializer, actual implementations may perform partial deserialization in a prior stage.

Many Serdes receivers also include a feature to assist with data alignment of the output. Most applications organize data into bytes or words (groups of bytes). For 8B/10B encoded applications, data is organized into 10-bit encoded symbols. The initialization of the clock divider in the deserializer is arbitrary, and the data received on the parallel data bus will have an arbitrary alignment that is unlikely to match the byte or symbol boundaries of the protocol. This can be corrected by downstream logic to steer data onto the appropriate byte, symbol, or word boundary. Alternatively, many Serdes receivers provide an input which forces the deserializer to "slip" one bit. Downstream logic detects

Marvell Semiconductor, Inc. - Ex. 1014, Page 000026

that data is not aligned to the appropriate boundary, and repeatedly pulses the deserializer control until the data "slips" to the desired alignment.

### 1.3.2 Equalizers

The interconnect between the transmitter and receiver device (known as the *channel* ) acts as a filter at typical baud rates, and distorts the serial data signal to varying extents. Figure 1.9 illustrates this distortion: The input waveform is a clean digital signal, but the output waveform is significantly distorted. The illustrated frequency response function for the channel is characteristic of a low-pass filter. Signal distortion occurs because the signal baud rate is above the cut-off frequency for this filter.

Signal integrity concerns frequently dictate that the data signal be equalized at the transmitter and/or receiver in order to counter the effects of the channel and decode the signal properly. Many variations on filter architectures are possible, all of which accomplish this. Fig. 1.10 illustrates the addition of an equalizer at the transmitter with a transfer function that is roughly the inverse of the channel's frequency response. This equalizer distorts the signal at the transmitter output such that the resulting signal at the receiver input is a clean waveform.

Most Serdes transmitter implementations include a FFE. The block diagram for a three-tap FFE is shown in Fig. 1.11. The serial data signal is delayed by several flip-flops which implement the *taps* for the filter. Each tap is multiplied by a *tap weight* value (also called a *filter coefficient*), and the results are summed and driven to the serial data output. FFE operation is described further in Sect. 3.2.1.

The number of FFE taps on the filter, the spacing of these taps relative to the baud rate, and the granularity of these tap weight values vary based on implementation. The terminology *preemphasis* or *deemphasis* refer to the FFE architecture, and indicate whether the data signal amplitude is increased or decreased as compared to the nonemphasized value by the FFE tap. The terminology *precursor taps* and *postcursor taps* refer to whether the FFE filter taps operate on an advanced or delayed signal (respectively) relative to the $t = 0$ tap. *Baud-spaced* taps are defined as taps where the delay from one filter tap to the adjacent tap is one-bit time interval; fractional spacing of the taps is also possible.

The FFE tap weights are selected to generate a filter with the inverse transfer function of the channel transfer function. Various algorithms exist for determining optimal FFE coefficient values; some select filter coefficients to maximize signal amplitude at the receiver, while others optimize eye width (i.e., minimizing jitter). More complex algorithms may search for an optimal trade-off between amplitude and jitter in order to optimize a more complex parameter (such as projected BER). FFE tap weights are determined for many applications by design and coded as fixed values within system software, however, there are some applications where FFE tap

Marvell Semiconductor, Inc. - Ex. 1014, Page 000027

weights are adjusted dynamically by the protocol based on signal characteristics at the receiver.

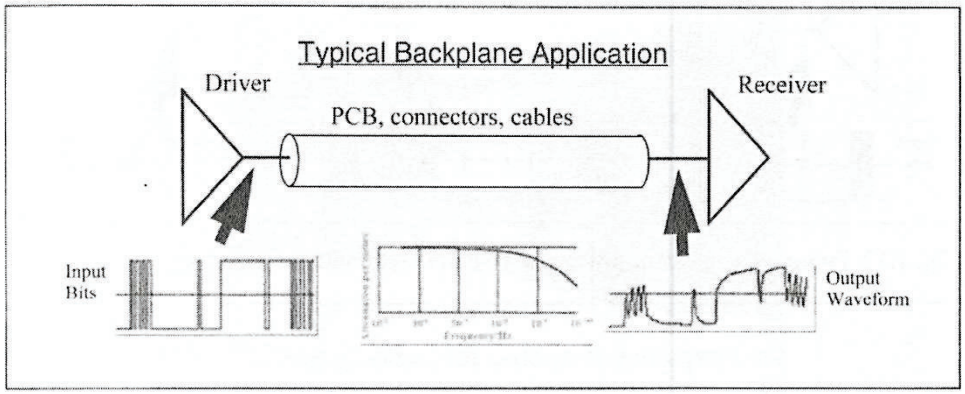



**Fig. 1.9** Signal distortion for a typical channel application

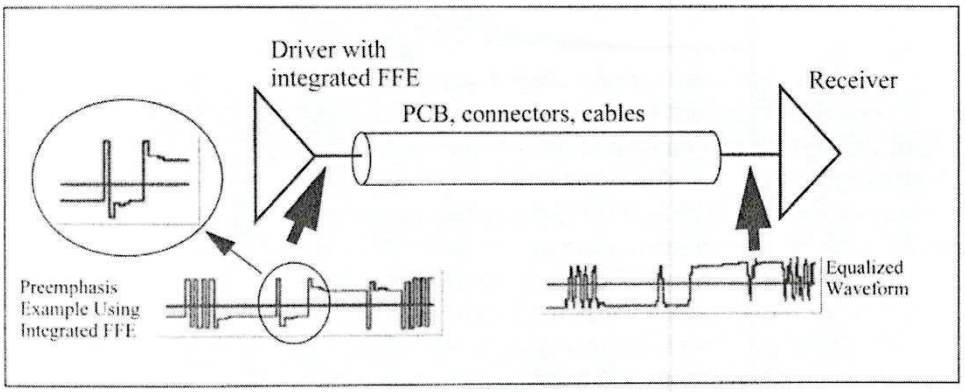



**Fig. 1.10** Typical channel application with equalization at the transmitter

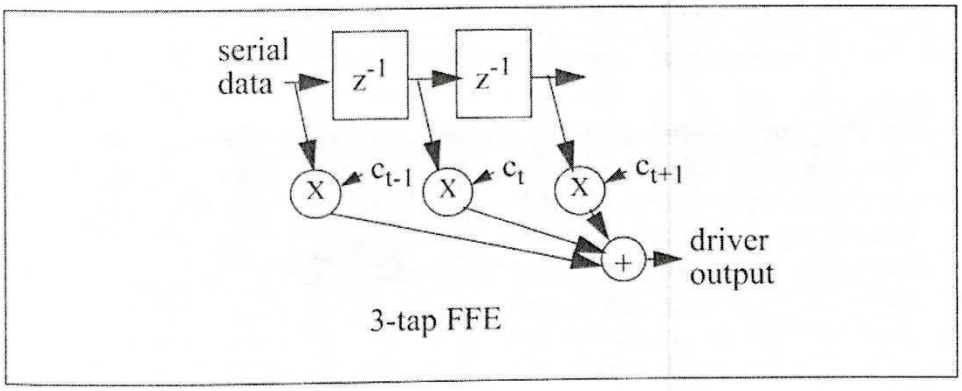



**Fig. 1.11** Three-tap feed forward equalizer operation

Marvell Semiconductor, Inc. - Ex. 1014, Page 000028

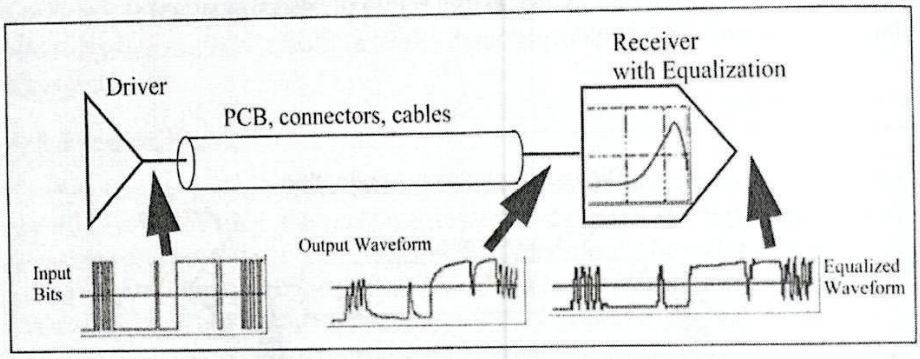



**Fig. 1.12** Typical channel application with equalization at the receiver

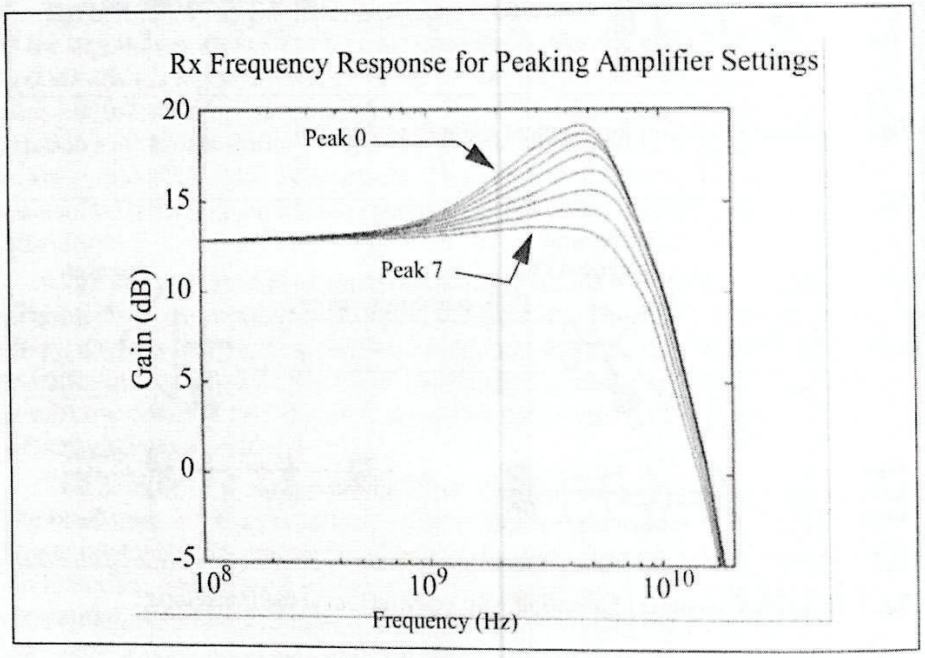



**Fig. 1.13** Receiver frequency response for peaking amplifier settings

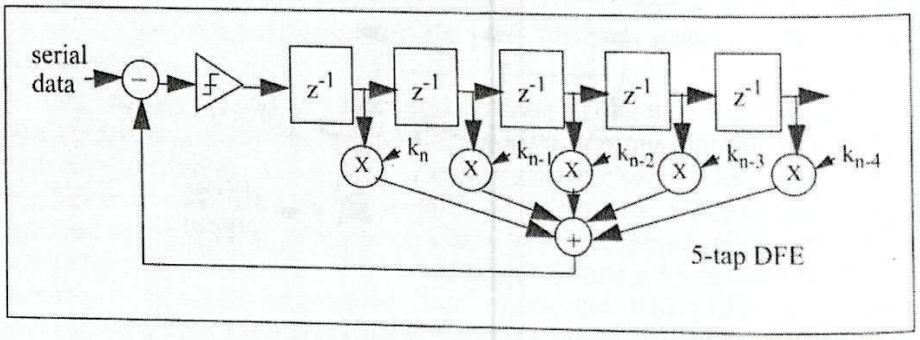



**Fig. 1.14** Decision feedback equalizer architecture

Marvell Semiconductor, Inc. - Ex. 1014, Page 000030

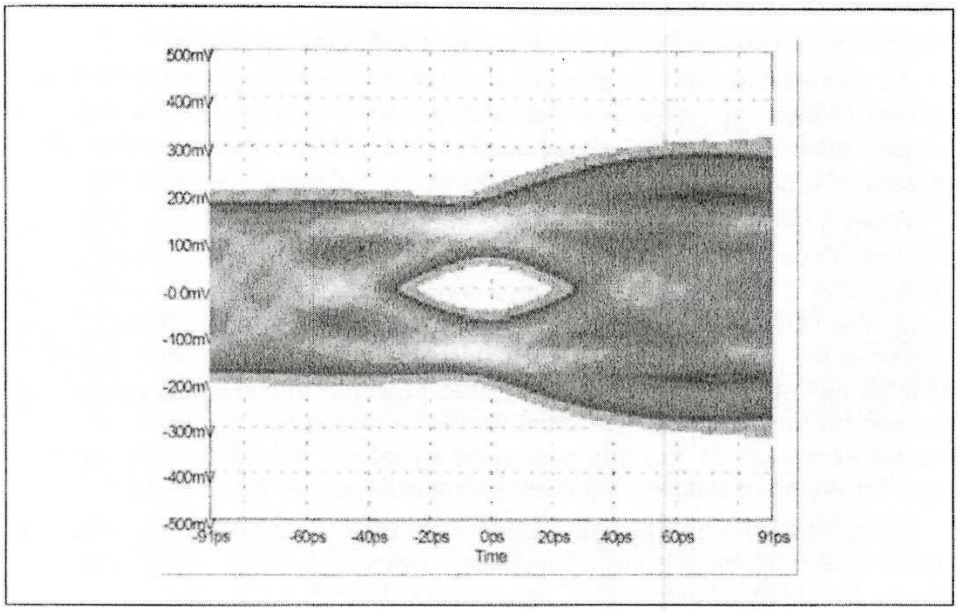



**Fig. 1.15** Virtual eye after equalization

Equalization may also be performed at the receiver as illustrated in Fig. 1.12. Despite the signal distortion at the input of the receiver, this equalizer corrects for the distortion and produces a clean waveform. For lower speed or lower loss links, the most prevalent approach is to use some variant of a peaking amplifier. Peaking amplifier circuits amplify the higher frequency signal components more than the lower frequency components. If the peaking amount is matched to the high frequency loss (difference between high frequency and low frequency), then the channel is equalized and the eye is opened up. Some Serdes devices allow programmable peaking levels; the frequency response of such a peaking amplifier for various provisioned settings is shown in Fig. 1.13.

For higher baud rates, the transfer function of the channel can cause jitter exceeding the bit width of the data and significant loss of signal amplitude at higher frequencies. A DFE stage is often included in receivers for these baud rates in order to recover data despite the otherwise "closed" eye.

A conceptual block diagram of a DFE circuit is shown in Fig. 1.14. The serial data signal is applied to a slicer circuit which makes decisions as to whether the incoming signal is a "0" or a "1". The received serial data is then delayed by a number of flip-flops which implement the filter taps. Each tap is multiplied by a corresponding tap weight value, and the results are summed. This sum is then used to correct the amplitude of the incoming signal, affecting the decisions made by the slicer circuit. Slicer decisions are thus affected by feedback based on prior data received. Although not shown in Fig. 1.14, some DFE architectures use feedback to affect both the amplitude and the sample

Marvell Semiconductor, Inc. - Ex. 1014, Page 000032

time of the slicer circuit. Such architectures adjust the CDR sample time from bit to bit based on feedback regarding the last several bits received.

As was the case for FFE circuits, the number of DFE taps on the filter, the spacing of these taps relative to the baud rate, and the granularity of these tap weight values vary. DFE implementations usually also contain logic which trains the DFE and sets DFE tap weights to optimal values dynamically.

Using a DFE, the closed eye shown in Fig. 1.5b is cleaned up to produce the virtual eye shown in Fig. 1.15. Note that the eye in this illustration is produced by a DFE architecture which corrects the CDR sample time from bit to bit. The DFE correction is valid for only one instance in time (based on the history of the previous bits). As such, once the DFE makes a decision as to whether the bit is "0" or "1", the DFE then proceeds to make adjustments for the next bit time which are different for the various signal traces of the composite waveform. For this reason, the signal eye shown in the figure is open for the bit of interest, but does not appear open for adjacent bits.

Many variations on equalizer architectures exist. As the baud rate increases, equalizer architectures become increasing complex. In some cases, protocol standards specify a base level of required equalizer functionality.

### *1.3.3 Clock and Data Recovery (CDR)*

Conceptually, CDR circuits monitor transitions of the data signal and select an optimal sampling phase for the data at the mid-point between edges. Since the timing of data transitions includes a jitter component, the CDR must perform some averaging to provide stability of this sampling point from one bit to the next. Intersymbol interference (ISI) and other components of deterministic jitter (DJ) are dependent on the spectral content of the data signal, and this frequency spectrum does change based on the data content. Shifts in this frequency spectrum sustained for hundreds of bits or more cause the CDR to adjust the optimal sampling phase dynamically.

CDR architecture is discussed further in Sect. 3.3.1. Features of the CDR may be of some significance to the Serdes user are discussed below.

#### 1.3.3.1 Maximum Run Length

A significant parameter for the Serdes which is primarily the result of the CDR design is the maximum number of consecutive "0" or "1" bits which can be received before the sampling point of the CDR risks incorrectly sampling the bits. An excessively long run of consecutive bits of the same value means that the CDR is not detecting any data transitions, and therefore cannot recover any clock information to ensure the data continues to be sampled in the center of the eye. A small drift in the sampling point relative to the baud rate of the data may cause the CDR to sample more "0" or more "1" bits than were actually transmitted. Also, the sampling point may require recentering when data transitions resume, and additional bits may be sampled incorrectly as this adjustment occurs. Some CDR implementations drive the receive data to a PLL and use the output of the PLL as the sample clock; clock outputs of the

Marvell Semiconductor, Inc. - Ex. 1014, Page 000033

receiver may change frequency or stop when such CDRs do not receive data transitions for a sustained length of time.

The maximum run length of consecutive "0" or "1" bits which must be tolerated depends on the protocol application and the data encoding defined for a given protocol. For example, protocols which use 8B/10B encoding are guaranteed to have no more than 5 bit times between data transitions. Protocols using another common encoding, 64B/66B, are guaranteed to see run lengths no longer than 66 bit times. Scrambled protocols may encounter much longer run lengths, and must determine requirements using statistical analysis. For example, Sonet/SDH is a scrambled protocol which specifies systems must meet a BER of $1 \times 10^{-12}$. It is generally accepted that run lengths of scrambled Sonet/SDH data longer than 80 bits statistically occur less frequently than the specified BER. Therefore, a Serdes used to receive Sonet/SDH data must tolerate a run length of 80 bits.

The run length which can be tolerated by a CDR design is related to the frequency tolerance between the two clock sources. In a system using plesiosynchronous clocks, the reference clock used by the receiver (and the CDR circuit) may be running at a slightly different frequency from the reference clock used by the transmitter, as is described further in Sect. 4.1.3.1. The frequency tolerance between the two clock sources is generally specified in parts per million (ppm). In a plesiosynchronous system, the CDR must continually correct the phase of the sample clock to remain in the center of the data eye. During periods where no data transitions are being received, the error in phase position builds up. Therefore, as the frequency tolerance of the system is increased (corresponding to larger allowed frequency difference between the clock sources), the run length which can be tolerated by the CDR design is reduced for a given performance (BER) target.

#### 1.3.3.2 Clock Operation During System Initialization

In the above discussion, it was noted that some CDR architectures derive the sample clock from the received data using a PLL. During system initialization or during system operation when cables are unplugged, etc., no data transitions are received for a substantial period of time. For some Serdes, this results in clock outputs of the receiver changing frequency or stopping. Any downstream logic clocked by these clock outputs must be designed to be tolerant of this frequency change or to assume logic is not clocked during these periods.

### *1.3.4 Differential Driver*

The differential driver stage is an analog circuit which drives the *true* and *complement* legs of the differential signal. Output data must be driven such that jitter is minimized. In some architectures, data is latched in a flip-flop clocked at the baud rate, and the output of this flop is driven onto the differential output. Such implementations require an internal high-speed clock running at the baud rate. This is illustrated in Fig. 1.16.

Marvell Semiconductor, Inc. - Ex. 1014, Page 000034

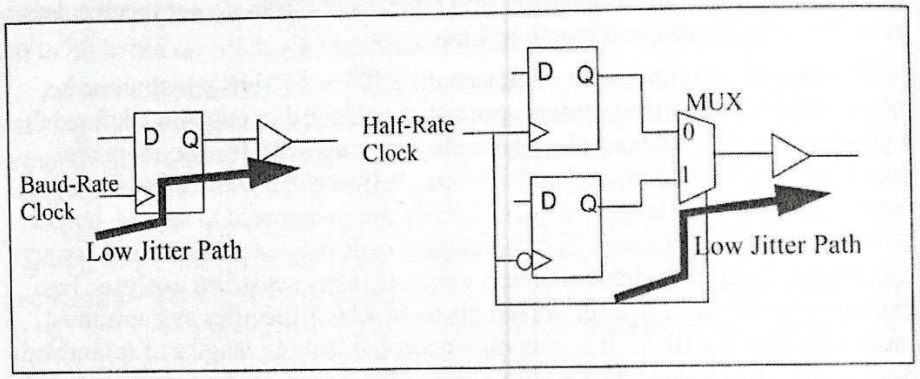



**Fig. 1.16** Driver stage architectures

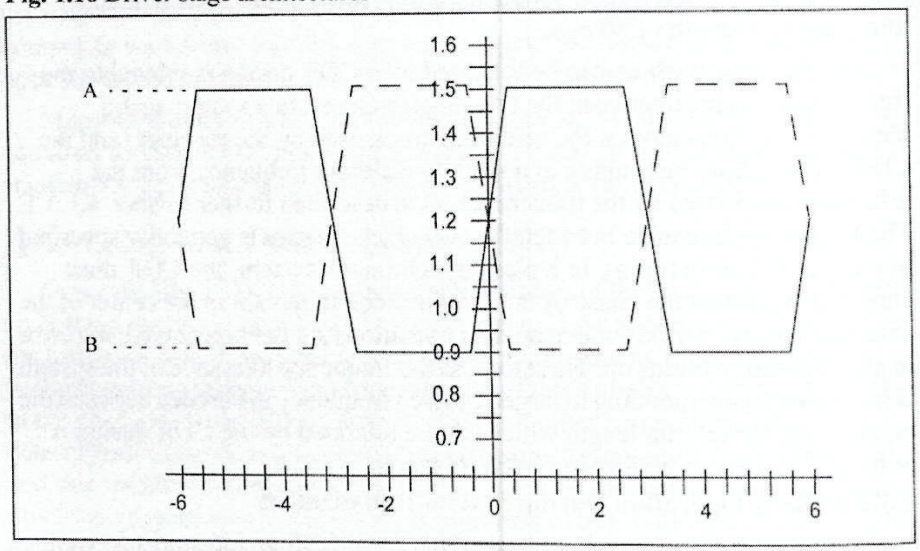



**Fig. 1.17** Single-ended complementary signals

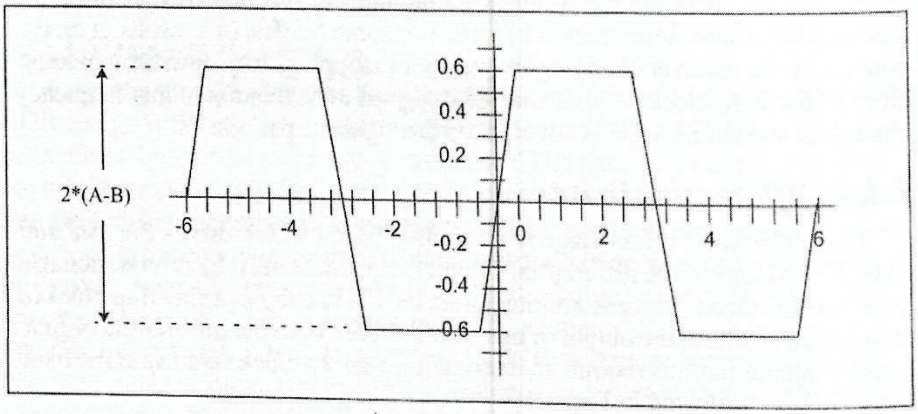



**Fig. 1.18** Differential peak-to-peak signal

Marvell Semiconductor, Inc. - Ex. 1014, Page 000036