

# TRENDS IN EXTREMELY HIGH SPEED DATA TRANSFER AND THE CHALLENGES THEY PRESENT

*Dr. Richard G. C. Williams*

Texas Instruments, Northampton, UK. Email: [richard@ti.com](mailto:richard@ti.com)

## ABSTRACT

As required data transfer rates increase, parallel interfaces linked to ASIC clock speeds become infeasible. Serial interfaces sending data at 3Gbps are common, at 10Gbps are available and at 25Gbps are being designed. Built from circuits that transmit and receive signals with bandwidths and switching speeds of many GHz using low voltage CMOS, these interfaces present exciting design challenges that stretch both the circuit designer and the silicon process technologist.

## 1. INTRODUCTION

The world is demanding faster processing of increasing amounts of data. Super computers are being used to predict weather patterns more accurately. The amount of traffic on the information super highway drives the need for huge routing systems that handle terabits of data per second. The proliferation of mobile phones and the extension of the services that they provide requires significantly more processing in the cellular base stations that distribute the information. The only way to satisfy this demand is to split the processing into smaller parts and to pass data, quickly and robustly, between the processing elements. There is a need for extremely high speed, virtually error free data transmission over relatively small distances. This is achieved by interfaces that turn internal parallel data into a serial stream before transmission and back to parallel data upon reception. Serializer/deserializer (SerDes) interfaces have enabled recent advances in many areas and are becoming an integral part of most systems. The key to their success is their small size, low power consumption and minimal error rates. In certain applications, such as memory access, they also need to have low latency. However, as data transfer rates increase to 10s of Gbps, maintaining these attributes becomes more difficult. Successfully solving the problems encountered will enhance current applications and open up new markets. The design of analogue and digital circuits that are included in SerDes modules pushes the state of the art. Reductions in the geometries of process technologies from 90nm to 45nm and lower provide help but architectural changes and extreme care in circuit design are needed in addition to this. As the speeds of component circuits exceed 5GHz, digital circuits start to be designed with analogue techniques and analogue circuits struggle for bandwidth, their performance beginning to be dominated by parasitics and mismatches.

## 2. THE PROMISE OF SERIAL LINKS

Any system that has a large amount of data to process in a short amount of time, such as a high speed router in a network, needs to split the work into several different processing steps. The amount of data that needs to be transferred around the system precludes the use of parallel links running at speeds related to the internal clocking of the ASICs. Serial links are used to reduce the pin count and the power consumed in these data transfers. Data will enter the system through a physical line module typically across several Ethernet or PCI Express links or through optical modules. This data will be channeled to one or more processing engines. There will be some amount of memory linked to the processing engine and in most cases the efficiency of storing and fetching data will affect the speed at which the data is processed greatly. Finally, this processing engine will be linked to other processors and other parts of the system typically across a backplane interconnecting several boards. All of these links must provide high speed data transfer and should be virtually error free with specified bit error rates normally being less than  $10^{-15}$ . The interfaces are split into short reach links (< 8" trace length possibly with 1 connector) and long reach links (< 50" trace length with several connectors). A typical but hypothetical system is shown in the figure

below. This configuration represents the state of the art in networking equipment where a single chip can have 360Gbps worth of I/O capability. This requires a large number of SerDes to be integrated on one piece of silicon.

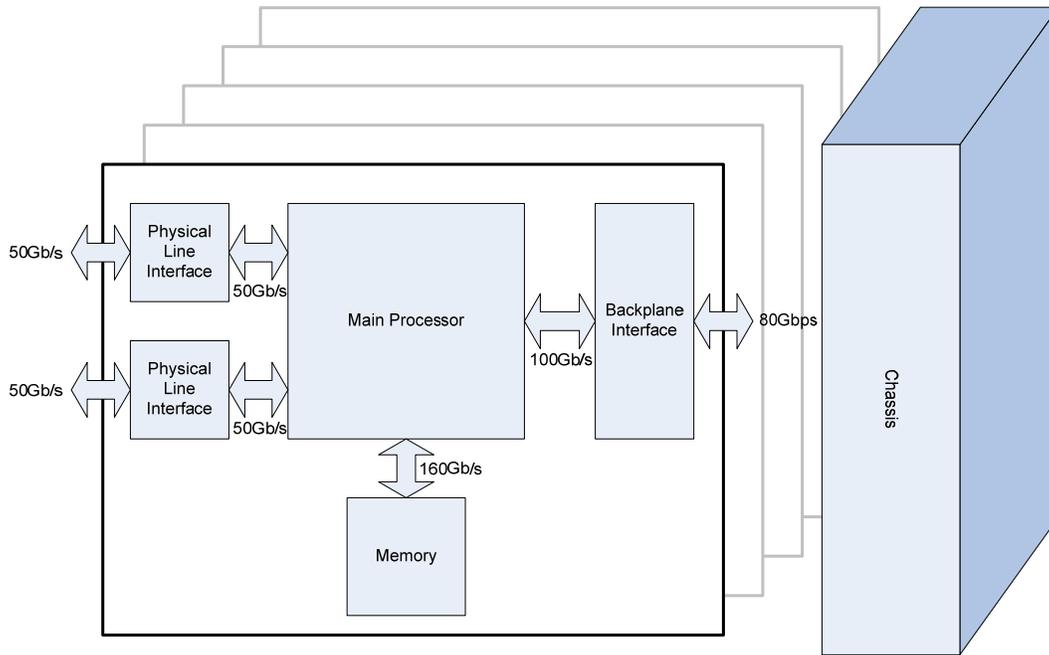


Figure 1: A typical system

Once a set of boards are designed and connected together in a chassis there is a need to upgrade the data handling ability of the system by only increasing the processing speed of the chips and the transfer speed of the interfaces. Increasing the data rate on a given link inevitably leads to an increase in power consumption and often in silicon area too. Going to a higher number of parallel input bits to be converted also tends to increase the latency of a link. All of these effects are undesirable. For economic reasons, the power supplies to the system and the cooling methods used must remain mostly unchanged. This limits the increase in power consumption that is acceptable. Similarly, there will be limited space for growth so any area increases in components must be minimized. An ideal situation would be doubling the data rate of an interface in the same area for the same power consumption. Obviously if area and power consumption scale in line with data rate there is little incentive to upgrade the system. In practice a combination of scaling in silicon technology and architectural changes allow improvements to be made as the table below illustrates.

Standard	Data rate (Gbps)	Bandwidth (GHz)	Coding, modulation	Rate over 64 pins (Gbps)	Approx power (W)
GbE	1.0	1.25/2	8b/10b, NRZ	16	N/A
XAUI	2.5	3.125/2	8b/10b, NRZ	40	2.0 @ 90nm
XFI	10.0	10.3125/2	64b/66b, NRZ	160	4.0 @ 65nm
N/A	20.0	20.625/4	64b/66b, PAM4	320	5.3 @ 45nm

Table 1: Trends in data rate and power for a fixed system with increasing serial link data rate

Unfortunately the data transfer problem gets more difficult as the data rate increases due to the physical characteristics of the communication channel. This limits the advances that can be made in a given system structure. We appear to be at the point where current systems are probably going through their final upgrades to data transfer speeds in the region of 10Gbps. Systems that need to transfer data faster than that

will need new architectures. Many systems are being redesigned to place clusters of chips closer to each other, minimizing the number of long range communication links that are used. There is a clear growth path for data transfer across short range communication links to reach and exceed 20Gbps rates in the next few years. Extremely short range interfaces using SerDes between pieces of silicon in the same package are also being considered. Once new architectures have been deployed it should be expected that the cycle of increasing data transfer rates will repeat.

### 3. SERIAL LINK REQUIREMENTS

Typical attenuation curves for the type of board traces that are used for data transfer are shown in Figure 2. It can clearly be seen that trace length and frequency have a significant effect on the received signal strength. Other factors such as discontinuities in connectors and vias that cause reflections add ripple to the frequency response of the channel. In order to receive a signal across these channels equalization techniques must be used. With a simple 2 level scheme like NRZ, the signal occupies an amount of bandwidth equal to half the signaling rate. Thus, at 1.25Gbps the channels look quite smooth and even in the most demanding links the attenuation can be removed by transmitter de-emphasis. As the data rate increases the amount of attenuation and ripple that needs to be compensated requires receive equalization techniques. Linear equalization is adequate up to data rates around 5Gbps. At higher transfer rates than this solutions across more than a few inches of trace length must turn to transmit de-emphasis techniques with multiple coefficients and receivers with a decision feedback equalizer (DFE). DFEs are particularly useful because they can counteract ripple. They also have the property that they only equalize the signal and so any interference from other sources, such as crosstalk in a system with many links, is not amplified by the equalization. As the number of links in a system increase all of the equalization techniques employed need to be adaptive in order to reduce the set-up time and optimization of link performance.

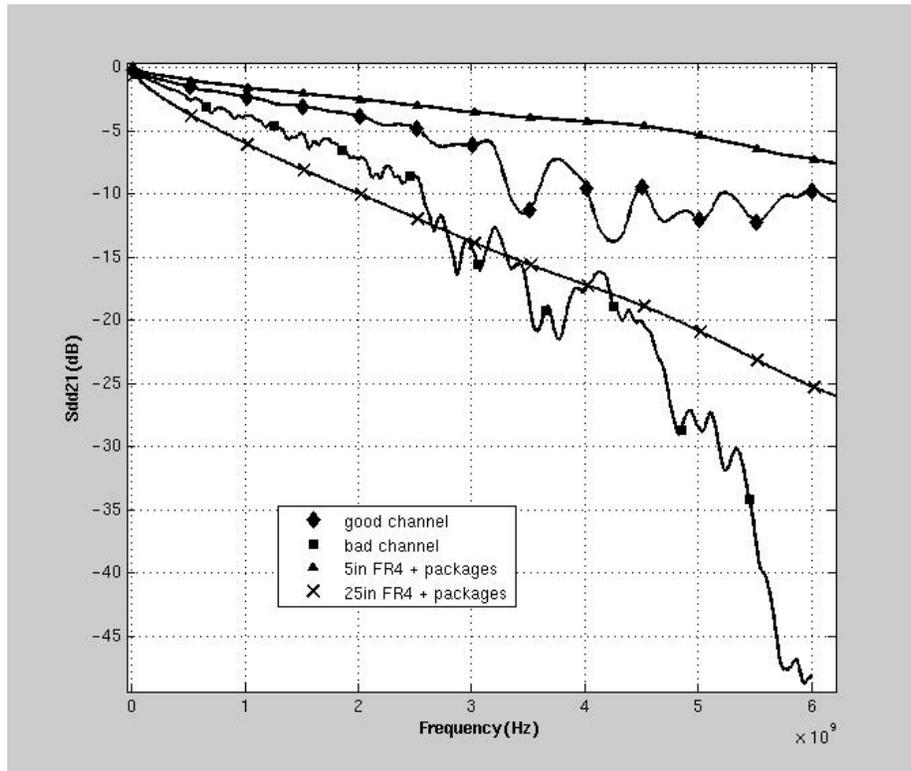


Figure 2: Examples of channels

Either an analogue or a digital approach can be taken to realize a receiver with a DFE. Analogue implementation is well known and has excellent coefficient resolution but it consumes a large amount of

power and area. Digital implementations have advantages for the future. Once size and power constraints have been met, the porting of the architecture to smaller process geometries becomes easier and yields further power and area advantages. Also, techniques such as feed forward equalization and test features become readily available and future digital signal processing extensions can be considered. As process geometries shrink it is expedient to use as many digital techniques as possible. Figure 3 shows a block diagram of the SerDes link designed with a digital DFE.

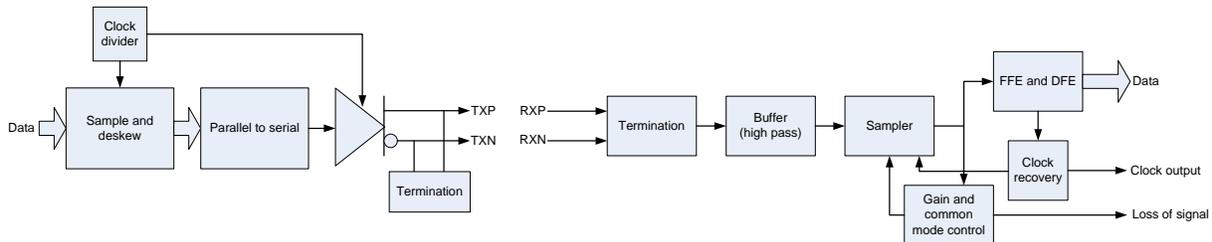


Figure 3: SerDes architecture including digital equalization

Running circuits at clock speeds equivalent to the data rate consumes exponentially increasing power as the data rate increases. Similarly, as the bandwidth that the signal occupies increases with data rate the amount of equalization needed to recover data with near zero bit error rate from a signal increases dramatically. Adding the extra circuitry to provide this equalization adds to the power consumption and area of the solution. Both of these effects eventually limit the data rate that it is economical to deploy on some systems. As the length of the link increases the receive signal strength falls significantly. Boosting such a weak signal becomes increasingly difficult with the falling operating voltages of the smaller process geometries that are needed to meet the goals for the area and power consumption of the SerDes. This ultimately limits the links across which higher data rates can be used.

## 4. DESIGN CHALLENGES

### 4.1 Process challenges

There are many difficulties encountered in the design of these extremely fast circuits using deep sub-micron CMOS technology. Challenges are present in both the digital and the analogue domains and at both the transmitter and the receiver. Many of these challenges are due to a change in underlying assumptions brought about by small process geometries. The first requirement of all designs is that they must be designed for variability. A pair of the same transistors placed next to each other on a die can behave differently and have different timing. Many circuit designs that rely on biasing circuits to behave in a consistent manner across process variations can no longer rely on these to the same extent. Extra steps need to be taken to determine how robust these assumptions are. Even digital circuits suffer from the variability in timing of the transistors. While this does not affect digital designs with many gates between registers, at high clock rates and few gates the variability is pronounced. Static timing analysis is being replaced by more statistical methods to try to account for this variability. The margins that are typically applied to ensure robust operation within the desired process, voltage and temperature ranges are no longer able to be applied without careful thought. Standard margin techniques lead to impossible design constraints. Due to this, a much more thorough understanding of the usage scenarios of the circuits is needed. Rules of thumb for jitter budgets, IR drop analysis and noise immunity have to be rethought and new analysis tools are needed. One method being used to reduce the spread of parameters that a design needs to cover is to vary the power supply in accordance with the strength of the silicon.

An area that SerDes are particularly sensitive to is jitter. Since the speed of the transitions in the signal is increasing there is less jitter margin before errors are made. The variability of transistors spreads out the jitter skews making the control of jitter extremely difficult. Exacerbating this are the transistor mismatches that create greater levels of common mode to differential conversion and lead to increased sensitivity to

power supply noise. Some of the standard methods of combating this, like extending gate lengths and widths, are no longer available in the very small geometry processes forcing designers to look for alternative solutions. In some cases, clock recovery becomes so problematic that the best solution is to revert to forwarded clocking schemes.

#### **4.2 Shrinking supply issues**

Another major problem is the reduction in supply voltage. The standard supply voltage at 65nm is 1V and this is likely to go lower as geometries shrink. With supply variations the designer must assume that the supply voltage is 900mV. The situation is made worse by the thinner metal stacks that are used leading to poorer power distribution and higher IR drops. In certain regions of a chip it may be prudent to design to a supply voltage as low as 850mV. It is not uncommon for the signal that a circuit must accept or produce to be above its supply voltage. In the receiver this leads to the need for gain control circuits. In the transmitter charge pumps or inductors are needed. Obviously standard inductors are too large for these applications so compact inductors must be designed. Inductors can also be used for bandwidth expansion to cancel the capacitance that loads high frequency circuits. Both active and passive inductive peaking is being considered. To allow circuits to maintain headroom, low  $V_T$  transistors are produced. These create problems with high leakage currents which consume too much power.

Back biasing is used to try to adaptively control the threshold voltage to make the trade-off between response time, headroom and leakage. Power management techniques are employed to turn off circuits when they are unused. Since this is addressing the problem of leakage power merely turning off the clocks to digital circuits is not sufficient. Investigations of the use of subthreshold design techniques are also underway although the speeds for these at the moment are an order of magnitude too slow.

#### **4.3 Digital compensation for analogue non-idealities**

The inevitable move to more digital content in SerDes means that samplers are replaced by structures resembling ADCs. The mismatches of the transistors in ADCs lead to another set of problems. Here INL and DNL effects start to destroy the required monotonic nature of the circuits. Offset compensation is necessary to keep these effects under control. Since these compensation circuits need to be switched in and out of the data path, extreme care must be taken in their design. Some compensation for the non-idealities of an ADC can be made in the digital domain. It is definitely in the designer's interest to make trade-offs between the complexities in both domains to come up with an efficient solution. Also, because the circuits are sampling at speeds close to the bandwidth of the transistors that they comprise of, the probability of a metastable event occurring increases. Such events will cause data errors which are not tolerable and so counter measures must be taken. Unfortunately these generally require more area and/or power which are contrary to the original design goals.

#### **4.4 Problems with going faster**

All of the circuits in a SerDes need to be faster as the data rate increases. In the case of digital circuits this leads to areas where standard clocked logic doesn't suffice. Targeted areas of the design need to consider dynamic logic techniques with custom cell design and manual placement. These areas of what is still considered digital logic need to be designed with analogue design techniques. Where clocked logic is used, the timing accuracy of the clocking needs to be carefully analyzed. This is also a problem in analogue circuits that contain clocked components.

All the analogue circuits need to be designed for extremely small response times. In several cases this leads to the need for extra circuits. For instance comparators need to be clocked, samplers need to be preceded by track and hold circuits, phase interpolators need to trade-off the difficulty of providing shrinking step size with performance needs and duty cycle correction circuits become essential. The constraints on clock generation also lead to new design choices. As ring PLLs start to push the limits of their bandwidth the jitter on the clocks they produce starts to be a limiting factor. LC PLLs can produce

cleaner clocks but currently occupy too much area. DLLs are an attractive option but it is difficult to design them at the speeds needed.

Just as in the case of falling supply voltages, bandwidth expansion techniques are also needed to push the speed of circuits. AGC circuits, analogue equalizers and track and hold circuits all need to operate with the signal before there is a chance to split it into parallel paths with reduced speed. Thus these circuits have to be designed for the full bandwidth of the signal. The efficiency with which the designer can provide high bandwidth in these circuits is one of the primary factors that will decide the trade-off between signal bandwidth and the number of signaling levels in the future.

## **5. CONCLUSIONS**

SerDes design presents an opportunity to solve cutting edge technical problems in an area where there are many uses for the technology. The proliferation of serial interfaces in large systems is well under way and new applications from on-chip interfaces to communications in mobile platforms are being developed. In all application spaces the only hindrance to wider adoption is solving interesting design problems. Problems from pure circuit design to silicon process issues are being worked on by a growing community of SerDes engineers. Taking these interfaces to 32nm technology and below, achieving speeds in excess of 25Gbps across long reach backplane links, creating a 50Gbps interface that consumes 100mW are all goals for the future. As soon as these goals and many others are achieved there will be applications waiting to deploy the solutions.

## **6. REFERENCES**

- [1] M. Horowitz, C.-K. K. Yang, S. Sidiropoulos, 'High-speed electrical signaling: overview and limitations', IEEE Micro, Jan 1998
- [2] J. L. Zerbe et al, 'Equalization and Clock Recovery for a 2.5-10Gb/s 2-PAM/4-PAM Backplane Transceiver Cell', IEEE JSSC, Dec 2003.
- [3] R. Payne et al, 'A 6.25Gb/s Binary Adaptive DFE with First Post-Cursor Tap Cancellation for Serial Backplane Communications', IEEE ISSCC, 2005.
- [4] M. Meghelli, 'A 10Gb/s 5-Tap-DFE/4-Tap-FFE Transceiver in 90nm CMOS', IEEE ISSCC, 2006.
- [5] J. Jaussi, 'A 20Gb/s Embedded Clock Transceiver in 90nm CMOS', IEEE ISSCC, 2006.
- [6] A. Varzaghani, C.-K. K. Yang, 'A 6GS/s, 4-bit receiver analog-to-digital converter with embedded DFE', Symposium on VLSI Circuits, 2005

## **7. ACKNOWLEDGEMENTS**

I have enjoyed learning about the issues in these applications from the engineers in the TI SerDes design community and from the customers that they serve. I am especially grateful to Andy Joy for his careful review of this paper and his helpful comments.