Novel new architectures for ultra-fast CMOS data converters in conjunction with complex DSPs are the key enabling components for optical transport systems at 100 Gbps and beyond. This paper reviews the fundamental CMOS technologies that have enabled real-world deployments of true single-wavelength 100 Gbps optical transport systems along with an analysis of the future possibilities for transport rates beyond 100 Gbps that will leverage continued advances in CMOS converter and DSP designs.

1. ICs for 100G systems

1.1. Challenges and successes for a single-chip 100 Gbps coherent receiver

To provide a long-haul 100 Gbps optical transport network with maximum reach and immunity to optical fiber non-idealities, the industry has settled on dual-polarization quadrature phase-shift keying (DP-QPSK) as a modulation method, which means that a coherent receiver is required. The biggest implementation challenge resulting from this decision is the need for low-power ultra-high-speed ADCs, and their technology requirements define the way that such a receiver can be implemented.

Without suitable ADCs – especially with low enough power consumption – it is impossible to produce a 100G coherent receiver suitable for a commercial optical network, as opposed to prototype systems only suitable for demonstration in the lab. In future such ADCs will be required for higher-speed short-haul links, where low power and cost become even more important because the number of these links is much larger than long-haul ones.

To fit within the power constraints imposed by the system, these ADCs need sampling rates of at least 56 GSa/s and resolution of 6 bits or more, with power consumption of no more than a few watts each. To accomplish this with sufficient dynamic range for input signals up to 15 GHz and higher, it was previously thought that this would require technology such as very advanced silicon-germanium (SiGe) or ultra-small-geometry CMOS (40 nm or smaller). At the end of 2008, by extrapolating from historic advances in ADC design, it was predicted that suitable ADCs would not be available until 2013 [1].

However, the development of new circuit techniques means that these ADCs actually became available in 2009 using 65 nm CMOS [2]. This brought forward the date at which single-chip 100G coherent receivers became technically and economically feasible, and has caused a significant change in the industry roadmap for these devices.

A DP-QPSK coherent receiver needs four ADC channels since there are two optical polarizations and each needs two ADCs to digitize an I/Q signal. To achieve 100 Gbps net line rate, a baud rate of at least 28 Gbaud/s is used to allow for overhead, which needs 56 GSa/s ADCs. The system SNR requirements mean that 6-bit resolution is typically required to allow some margin for added noise and distortion, so for four ADCs the output data rate to the DSP is 1.3 Tb/s, or 1.8 Tb/s if 8-bit resolution is used to allow more noise margin or digital automatic gain control (AGC) and equalization after the ADC.

If the ADCs are not integrated with the DSP then this huge amount of data has to be transmitted between chips, which is not only difficult to implement (very large number of channels with high data rate) but uses a lot of power for serialization and de-serialization as well as actually transmitting the data – even using an optimistic figure of 100 mW/ch for an 11 Gb/s channel (ADC transmit + DSP receive) means 3–4 W per ADC is needed just to transfer the data. This is acceptable for a prototype or demonstrator, but not for a production solution.

A 100G receiver DSP – which performs functions such as equalization, chromatic dispersion compensation, and data recovery – needs in the order of 50 M gates, which mandates the use...
of CMOS. The system power requirement for a complete coherent receiver is only a few tens of watts; since a 40G ADC DSP chip in 90 nm already dissipates more than 20 W [2], geometries of 65 nm or smaller, as well as power-efficient design techniques, are needed for a 100G receiver.

This implies that the ADC should also use CMOS although this means the design is extremely challenging. A single-chip solution is really the only viable way forward, especially in order to take advantage of future CMOS technology improvements – though this does assume that the ADC performance scales similarly to digital circuits, which may not be true for conventional ADCs. Even if a multi-chip solution can be built – for example, using SiGe ADCs together with a CMOS DSP in a multichip module (MCM) – then not only will the overall power be higher, but the production cost will be greater and the yield of such a complex solution will inevitably be lower. This also does not give a good roadmap towards even lower power and cost solutions for short-haul and beyond 100G.

Designing a 56 GSa/s 6–8b ADC in any technology presents major difficulties. The challenge is even more difficult in this case, because available power for the ADC + DSP is limited both by supply capability and thermal dissipation. A reasonable target is 10 W or less for a complete 4-channel ADC, which means little more than 2 W per ADC cell. Conventional ADCs with this level of performance dissipate much more power than this, due to the high power needed for the wideband sampling/demultiplexing front-end and clock circuits as well as the back-end ADC power.

To achieve such high speed and resolution multiple lower-rate interleaved ADCs are used, driven by one or more wideband sample-and-hold (S/H) circuits, usually with more demultiplexing in between. The S/H circuits need very wide bandwidth and low distortion, which is why either SiGe or very small geometry CMOS (40 nm or smaller) is normally said to be needed for 56 GSa/s. However, with the sub-1 V supply voltages imposed by 40 nm CMOS it is very difficult to design a S/H circuit with reasonably large signal swing (to preserve SNR) and linearity (to preserve THD); higher-voltage SiGe avoids these problems, but at the cost of much higher power dissipation.

The very small device sizes available in modern CMOS processes do combine low power consumption and high density with high speed, but this comes at the price of increased noise and mismatch. The normal solution to this is to increase transistor sizes (gate length and/or width), but this is not possible here because reducing bandwidth or increasing power consumption is not viable for the system application. Small transistors mean poor matching in both S/H and ADC, not just in the signal paths but also in the clock paths, where 100 fs clock skew causes —40 dBc distortion for a 16 GHz input signal. The only feasible way to reduce these mismatch-induced errors is by widespread on-chip calibration.

Providing clock skew adjustment is not so difficult in theory. However, measuring and calibrating skew down to sub-picosecond accuracy is a much bigger problem, especially while maintaining this accuracy over time and environmental variations without either being able to take the ADC off-line for calibration or needing a huge amount of complex data analysis (such as FFTs or correlation) to calculate the errors.

To overcome these challenges, one solution is to use a new sampler/demultiplexer architecture [3] which gives the linearity, noise and bandwidth required without needing extremely short-channel (40 nm or below) transistors; allows simple calibration of amplitude and timing errors during operation; and dissipates <0.5 W. Instead of a conventional S/H using analogue switches and sampling capacitors, this circuit generates controlled-shape constant-area (charge) sampling pulses which are then demultiplexed to drive a large array of 8b SAR ADCs (320 x 175 Ms/s) (Fig. 1).

Using SAR ADCs instead of full-flash means that increasing resolution from 6b to 8b carries only a small penalty in power and area. Moreover, the increased resolution reduces quantization noise and allows more margin for other noise contributions. It also opens up the possibility of doing some AGC digitally after the ADC instead of in the optical front-end, which has the advantage of perfect channel matching even with rapid gain changes to track optical power variations. Some of the extra SNR can also be used for digital equalization to correct for response errors in the ADC and package and PCB losses. SAR ADCs also scale very well with smaller technology since most of the power is digital.

The high circuit density of CMOS is used to good effect by having a large number of calibration DACs (more than 400 per ADC) to trim out all significant device mismatches in the signal path, including timing errors (skew) as well as amplitude errors (gain and offset). This circuit works in such a way that all these errors can be calculated in real-time in the background by simple analysis of all the ADC digital output data. This analysis is done inside the ADC, and the results are then read out at a much lower rate (microseconds to tens of milliseconds) and used to drive a low-complexity convergence algorithm with a time constant of typically less

---

**Fig. 1. Block diagram for 56 GSa/s ADC.**
than a second, which provides convergence at power-up and also continuously during operation to track any parameter drift.

Realizing the DSP for a 100G coherent receiver is hardly trivial, not just because of the extremely high processing power required (12 TOPS were needed [2] for a 40G receiver) but also the very wide data paths (thousands of bits wide) needed to keep clock rates reasonably low. The latter is the key to improving power efficiency, because high clock rates need more latches to reduce logic depth, which leads to increased power consumption. For optimum power consumption – the most critical parameter for a 100G receiver – the optimum clock rate for the DSP will be much lower than the typical CPU clock rates in a given process and will usually be around 500 MHz even in 40 nm and 28 nm processes. The silicon area will be larger than a faster-clock DSP but this is secondary to power consumption.

The DSP design problem then really becomes one of interconnect, not gate count. Unfortunately design tools (and designers) tend to design the circuits first and just accept the interconnect that results, especially when synthesizing the HDL logic from a high-level language description. This should really be done the other way around – by thinking of the signal flow and interconnect first, not the gates – otherwise the designer can end up with a DSP design which is difficult (or even impossible) to lay out. An architecture which seems optimum for pre-layout power and gate count can turn out to be far from optimum after layout and interconnect; this should not come as a surprise given that >90% of power is due to interconnect capacitance, not gate capacitance.

The unpleasant implication of all this is that it is possible to define an architecture and circuits and go all the way to synthesis and layout, only to find out that the resulting power is much too high and that this was not the right way to do the design in the first place. To avoid this tools are needed which can either go around the architecture-to-layout loop quickly, or can predict the interconnect consequences at the architecture stage. Unfortunately these tools do not really exist right now.

1.2. Advanced package design for 100G

Package design for a single-chip 100G device poses many electrical and thermal problems and standard design techniques and materials may not be enough to meet the severe requirements. Even assuming it is possible to design the ADC and DSP separately, there are other major problems to be overcome when integrating the two in a single-chip 100G coherent receiver: noise coupling between digital and analogue is one obvious issue, and by looking at approximate numbers the sheer scale of the problem becomes apparent.

The system designer now has a very large DSP with peak switching currents of $\sim 100 A$ (and perhaps $\sim 100 mV$ supply noise) on the same die as ADCs which need to have jitter of the order of 100 fs. Assuming a reasonable supply noise delay sensitivity of 1 ps/mV for the clock path, the analogue supply noise needs to be $\sim 0.1 mV$, which means $\sim 60 dB$ isolation between analogue and digital regions.

This might not seem so difficult until the bandwidth of this noise coupling is taken into consideration. This is several GHz or even tens of GHz because, unlike typical analogue or RF circuits, there is nothing to band-limit the noise. At these frequencies, noise isolation methods such as triple-well are no longer effective. In fact, given the frequencies and bandwidths involved and the jitter requirements, the noise coupling problem is probably more difficult than on any previously realized mixed-signal chip. However, it can be solved by very careful design, some new noise isolation techniques, and maybe a touch of “black magic” [4].

It is obviously undesirable to allow large DSP current spikes to get out of the package and into the PCB, so as to ensure good EMC performance and to prevent interference with the analogue circuits. Relying on PCB decoupling also gives the end user a perfect opportunity to get this wrong and degrade the performance in an unpredictable manner.

Ultra-low-inductance decoupling inside the customized flip-chip package and on the chip solves this problem, but any resonances between the internal L and C are then very high Q and always fall in the hundreds of MHz region, which is similar to the clock frequencies used by the digital circuits. So just adding decoupling on the “more is better” principle can do more harm than good if it makes a resonance coincide with a clock frequency. Simulations of the entire chip and package together are needed, and it will usually be necessary to take steps to damp the power-supply resonances such as damped on-chip decoupling and/or controlled-ESR ceramic decouplers in the package.

Once the digital noise has been kept under control it needs to be kept out of the sensitive analogue circuits. Noise coupling paths which have to be considered include not only the silicon substrate, but also the package substrate, underfill, die attach and lid – any of these can cause jitter which is synchronous with the DSP clocks, and is modulated by the time-varying DSP current.

With careful attention to all this it is possible to realize devices with very small performance degradation due to digital noise coupling – for example, as little as 0.05 ENOB (0.3 dB) difference between digital core reset and active has been measured on a single-chip coherent receiver as shown in Fig. 2.

Getting the ADC/DAC signals from the PCB to the chip is also difficult, not just because of the wide bandwidth and tight S11 requirements (to prevent reflections) but also because the package chosen for this type of application is typically a relatively largelysize- chip BGA (to fit in the chip and decoupling) instead of a small RF-style package. This means that very careful attention is needed to get low loss in the package and clean 50 ohm transitions from PCB-package and package-chip, while still being compatible with relatively standard multilayer PCBs (thick ceramic PCBs are not an option in these types of telecoms applications). Design techniques like using coaxial via transitions into chip bumps and dual 100-ohm lines for shorter via lengths all contribute to better signal integrity between chip die, package and PCB. For a truly system-compatible device, the package/SoC designer must also optimize the signal launch into a coplanar waveguide which is typically used to connect to the optical front end.

There are also unusually difficult thermal problems to solve; the power dissipation of these chips may not seem high compared to high-end CPUs, but the available cooling in the system is much poorer due to both space restrictions and high incoming air temperature. Every few degrees temperature rise on the chip makes it considerably harder to design a workable cooling system, even with best possible design practice such as using flat heat pipe heat spreaders.

This means that the required thermal resistance of the package from die to lid (to attach to heat sink) is much less than might be expected; if a die temperature no more than 10 °C hotter than the lid is specified, then a $\theta_{JC}$ value of around 0.2 °C/W is needed. To achieve this, together with high reliability, when mounted on a PCB may need specialist package materials such as an aluminium–silicon–carbide lid and indium–silver TIM (thermal interface material), as well as low-loss Hi-CTE ceramic substrates. These are not widely used right now in the ASIC marketplace, though they are well-established in applications like high-end server CPUs [5].

Ultimately, the correct choice of package and package design techniques has a major impact on the viability of the coherent receiver SoC. Performance factors such as low loss, better impedance matching (i.e., lower return loss) and high routing density, along with good thermal performance are key considerations.
1.3. Moving to single-chip 100 Gbps transceivers: greater integration and more functionality

For first-generation 100 Gbps line cards, the challenge is to integrate the electrical and photonic components on a PCB within the standard constraints of power and area. Those system designers following the OIF 100G MSA [6] for encasing the design in a module form factor typically face even more stringent constraints for these parameters. The trend in optical networking form factors is towards continual reduction in size. This is certainly clear when looking at the evolution of 10 Gbps transponders. Now that the 40 Gbps optical transport market is hitting mass market status with regards to deployments, it is likely to follow a similar trend in form factor reduction going forward, particularly when 40 Gbps links become more commonplace in medium-reach metro networks.

A similar trend is expected for 100G modules which may even overtake 40G deployments in medium-haul networks and router links. This may be driven in large part by the more cohesive industry effort to standardize for 100 Gbps, coupled with the strong demand for greater network bandwidth. In short, with new network application spaces and smaller form factors required for second and third generation 100G systems designs, anywhere the system designer can look to maximize integration by having fewer components and/or lower overall power and area is highly desirable.

Early 100G systems enabling true single-wavelength transport of a 100 Gbps signal are the first step towards mass deployment. These systems can take full advantage of a single-die CMOS coherent receiver chip including ADCs of sufficient speed and linearity and low power. A component level overview of an early generation system based on such a device is illustrated in Fig. 3 where the different sections of the coherent receiver IC are color-coded by function with the high speed analogue interfaces in green. The other primary components of this system include; an OTU-4 Frame/FEC device, a transmit multiplexer device (for QPSK encoding), laser driver/amplifier IC, transmit modulator photonics and optical receiver/hybrid device. In this instance, the DSP is only performing receive functionalities such as chromatic dispersion (CD) and polarization mode dispersion (PMD) compensation, demodulation, and equalization.

More robust performance for long haul links is another requirement driving the need for greater integration as the DSP has to take on yet more functionality. One example is Soft-Decision Forward Error Correction (SD-FEC), where the raw data output from the receiver to the error-correcting decoder has more than 1 bit to indicate probability of a 1 or 0 instead of just a binary hard decision. This decreases the error rate and increases the range of a system.

For 100G devices incorporating SD-FEC the complexity and size of the DSP is larger. The SD-FEC decoder also really needs to be on the same chip as the ADC/DSP because the data rate between them is increased to carry the soft-decision data. The higher line rate required for greater FEC overhead also requires a multi-channel ADC with a higher sampling rate (up to 65 GSa/s) in the receive path. For the physical implementation in CMOS, all this means moving to a smaller process node to achieve the higher gate count while still maintaining the power budget for the system.

Greater integration and lower power can also be achieved by including some of the transmit path functionalities in the DSP; for example, a soft-FEC encoder has much higher power dissipation when realized in an FPGA instead of in custom logic. An overview of this type of system implementation is shown in Fig. 4. Components are much the same as the first generation implementation in Fig. 3, but the DSP-based IC is considered to be a transceiver and not just a coherent receiver device.
Even greater integration can be realized by interfacing a high-speed CMOS DAC to the DSP to form the analogue output stage of the line side transmit path on the transceiver IC. This replaces not just the external high-power multiplexer (usually SiGe) but also removes the multichannel SERDES needed to drive it, thus saving even more power. An example of a system using this approach is shown in Fig. 5.

Here a high speed 4-channel 8-bit DAC (up to 65 GSa/s) is integrated on the same CMOS die as the DSP, along with the corresponding 4-channel ADC and high-speed 11 Gbps interfaces to connect to an OTU-4 Framer on the client side. The DSP transceiver IC is directly connected to the driver IC, eliminating the requirement for a separate transmit multiplexer device. A single-chip CMOS transceiver like this provides the lowest power and cost solution for 100G optical transport, with a clear migration path to higher-volume lower-cost applications in the future.

In addition to offering a more integrated solution, incorporation of a high-speed DAC allows for greater functionality in managing the transmitted signal to the fiber. A digital filter before the DAC can be used for waveform shaping and to compensate for frequency response errors, not just the DAC but also in the package and PCB. These include short-term reflections as well as losses and aberrations in the laser driver amplifier. It is also possible to add nonlinear predistortion to compensate for driver and modulator nonlinearity, which is essential if multilevel signaling is used to increase bit rates beyond 100 Gbps.

Utilizing a multi-channel DAC also allows much better management of skew between the four signal paths. This can now be done by adjusting the digital filter coefficients to give a precisely known and stable delay for each DAC channel which is identical for every chip and does not vary with voltage, time or temperature.

The transmit path is now the dual of the receive path. It uses four precise time-synchronous high-resolution data converters, with time, frequency and amplitude errors in both the converters and the external components compensated for digitally on-chip. This is perfectly suited to future shrinks in process geometry where the digital power and area fall even further.

2. Looking ahead to 400 Gbps/1 Tbps

With continued advances in high performance CMOS converters, DSP techniques and package design, system designers will be able to push the boundaries for realizing higher optical transport speeds in real-world systems. However it is clear is that the limit to what can be achieved at a particular process node is no longer the data converters or DSP complexity as such, it is digital power consumption. It is no longer possible to use all the available silicon...
area and keep within power budgets. Trying to squeeze closer to the Shannon limit in future by using ever more complex algorithms makes this problem even worse – such systems may work in the laboratory but be essentially unusable in the real world.

This is the same problem being faced by CPU designers, but made worse by more severe thermal limitations and the need to keep everything on one chip to minimize inter-chip data traffic.

Since the power reduction for each new process node is decreasing due to factors such as leakage and non-scaled supply voltages, keeping up with the projected bandwidth explosion will mean finding either new more power-efficient digital circuit techniques (e.g., adiabatic logic), more efficient algorithms, or some fundamentally new process technology (e.g., graphene). Without such advances, data rates in future will not increase fast enough to keep up with the rapidly increasing demand for data bandwidth. This is assuming, of course, that the issue of available optical bandwidth is addressed since the Shannon limit is already being approached.

2.1. Encoding options for long haul transport

As the industry moves from 40G/100G towards 400G system development, the challenges in dealing with fiber channel distortions will be even greater. A perusal of papers submitted to various technical conferences over the last couple of years shows that a broad base of research is ongoing in both commercial and academic sectors to investigate and experiment with options for optical transmission at line rates beyond 100 Gbps. One of the successful driving factors for 100G deployment has been the industry convergence and cooperation on a common encoding format: DP-QPSK. For the next generation transport rates, the options are varied and converging on an ‘ideal’ format will represent trade-offs among different parameters.

These include line rate, channel spacing, spectral efficiency, ADC sampling rate and ENOB. Choice of encoding format will also be affected by the decision on whether to keep to the existing 50 GHz optical grid spacing, or use wider bandwidth or even flexible bandwidth channels, and there is no clear consensus in the industry about this. Comparing DP-QPSK to larger constellation single-carrier encoding options for an optical transmission system with a line rate of 448 Gbps, some of the parameters are summarized in Fig. 6.

The trade-off for a comparatively simple encoding constellation (e.g., DP-QPSK) is that an extremely high ADC sampling rate is required (224 GSa/s). The on-chip digital interface between the ADCs and the DSP would be almost 7.2 Tbps (for 8-bit ADCs), which will pose a very difficult DSP interconnect design challenge. For a larger constellation encoding format like 256QAM a lower sampling rate is required, but the complexity and gate count of the DSP will be higher posing many challenges for the IC designer, not least of which is maintaining system power budget.

Another important consideration is that increasing the sampling rate and channel bandwidth will place even more stringent requirements on low-loss PCB and package technology, and indeed the entire path all the way from the chip to the optical modulator, which is challenging even now at 56–65 GSa/s. Though digital equalization can be used to a certain extent to compensate for these losses in a system using ADCs and DACs, a lot of equalization will reduce SNR too much, so the loss will have to be decreased further if sample rates rise. Unless the trace lengths can be reduced – difficult given thermal and physical constraints – this implies that lower loss materials will be needed, which are also compatible with high-reliability multi-layer PCB assemblies and lead-free soldering.

Improving the performance of CMOS ADCs and DACs to meet the requirements of 400G or even 1 Tbps – either by increased ENOB (e.g., to 7) or higher sampling rates (e.g., to 112 GSa/s or above) – is already known to be feasible, but it is likely that 20 nm technology or smaller will be needed to keep the DSP power consumption under control at such high data rates. Since it may also be necessary to reduce further or completely remove ESD protection circuits, if bandwidths double, this adds to the difficulty since such small geometry CMOS is even more ESD sensitive in the first place.

Another candidate for encoding formats for transport rates beyond 100 Gbps is OFDM (Orthogonal Frequency Division Multiplexing) which is a multi-carrier approach to gaining greater spectral efficiency from the transmission channel. Like ADSL, one advantage of OFDM is that it can allocate varying bit loading across the channel to obtain the maximum bit rate that the channel is capable of, so better channels can automatically carry more data. The downside of this is that the network has to deal with channels of different – and possibly even time-varying – bit rates, which could be extremely challenging for network management and may mean that a major change in network architectures is needed.

3. Optimization for long-haul versus medium and short-reach links at 100 Gbps and beyond

Much of the industry focus for 100 Gbps transport, both in research and in commercial network outlay, has been on long-haul links in the core network purely to cope with increasing capacity demand. This is no different from historical network growth patterns for 40 Gbps and 10 Gbps but the proliferation of 100G

<table>
<thead>
<tr>
<th>Modulation Format</th>
<th>DP-QPSK (4-QAM)</th>
<th>DP-16QAM</th>
<th>DP-64QAM</th>
<th>DP-256QAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line rate</td>
<td>112 Gbaud</td>
<td>56 Gbaud</td>
<td>37 Gbaud</td>
<td>29 Gbaud</td>
</tr>
<tr>
<td>Spectral efficiency</td>
<td>4 b/s/Hz</td>
<td>8 b/s/Hz</td>
<td>12 b/s/Hz</td>
<td>16 b/s/Hz</td>
</tr>
<tr>
<td>OSNR (BER = 10^-3)</td>
<td>19.3 dB</td>
<td>23.1 dB</td>
<td>27.3 dB</td>
<td>31.9 dB</td>
</tr>
<tr>
<td>Sampling rate</td>
<td>224 GSa/s</td>
<td>112 GSa/s</td>
<td>74 GSa/s</td>
<td>56 GSa/s</td>
</tr>
<tr>
<td>Effective no. of bits (ENOB)</td>
<td>3.8</td>
<td>4.9</td>
<td>5.7</td>
<td>7.0</td>
</tr>
</tbody>
</table>

Fig. 6. Encoding options for 448 Gbps [7].
ports is expected to move through the network faster due the fact that the rate of increasing capacity demand is itself increasing. The demand has been fuelled by the advent of suitable CMOS technologies for coherent detection and the industry’s strong efforts to standardize 100G transport technologies.

Migration of 100 Gbps links to medium (e.g., Metro spans) and even short-reach distances (e.g., intra-office/intra-datacenter) could be seen in the not-so-distant future. As coherent detection-based solutions mature and become more ubiquitous, markets for these shorter distances may be a logical next step. While complicated FEC and dispersion compensation may not be as critical because the distances are shorter, the demand for sheer volume of data will eventually push towards a coherent-detection approach, in combination with a multi-level or multi-phase encoding format, to maximize throughput and reach. This is especially true if cost-down approaches require using the same optics for lower rate transport rates (e.g., 100 Gbps over 10 Gbps optics).

Though a single-chip CMOS transceiver can certainly be produced at a suitable price in large enough volumes, there will have to be a step change in both the cost and production capacity of suitable optical modules away from the almost hand-built ones used now for coherent 100G systems. This may be a bigger challenge still to the industry.

Even at the very short reach interconnect level (chip to chip, chip to module) consideration of new techniques for high speed interfaces will have to come into play. Once system requirements dictate the need for single interfaces transporting data rates greater than 28 Gbps, traditional approaches to designing these interfaces will hit a wall. This is no different from what happened with the progression from POTS modems to ADSL, and from 10G coherent systems in long-haul. When a transmission channel capacity limit is reached (and no more channels can be built), the only option left is to increase modulation complexity and push towards the Shannon limit for that channel.

In both these cases – short-haul optical and electrical – the problem is not just how much data can be pushed through the channel, but how to do it at the lowest cost per bit. Right now the numbers favor simple binary systems with sub-optimal capacity, but this will not always be the case – and when this alters there is likely to be a seismic shift in the optical industry.

4. Configurable transceivers

Looking to the future, it is also conceivable that the next step towards greater functionality and flexibility in optical transceiver IC design will include a move to configurable or “software-defined” transceivers. The driver here will be to build systems that have greater channel and network “awareness”, and which can cope with a wider range of impairments, data rates and encoding formats. To enable such a high degree of flexibility will require very complex DSP designs that may likely not be feasible to implement in real-world systems until process geometries make them possible.

It may never be possible to afford the power increase needed for such flexibility either. Even the most efficient reconfigurable logic solutions carry at least a 2× overhead over custom hardwired logic, which although still much better than the 10× typical for FPGAs, may be difficult to justify if algorithm complexity continues to increase.

On the other hand, a reconfigurable system means that algorithms and firmware can be changed after chips have been manufactured, which not only allows unforeseen changes to be made but also fixes to bugs found in the field. If this also speeds up the design and verification time because changes can be made later, it is possible that reconfigurable chips could get to market faster and in a more advanced process to make up for the power penalty.

5. Silicon process nodes for 100 Gbps and beyond – is smaller always better?

As this paper is written from the point of view of IC designers, it would be useful to take a look at design choices with respect to base process technology. A common view in the technology industry is to target the latest technology, which in the case of IC design is the most recent CMOS process. The rationale is lower power, more functionality in a given silicon area, or reduced silicon area for the same functionality. These criteria are typically applicable to high-volume designs such as CPU or FPGA where working chips can be graded and sold at different price points based on performance, or consumer products (especially mobile or PC) where low unit cost and high integration levels are necessary to succeed in the market.

In both cases, there is a strong drive to maximize integration levels and functions on a chip, to achieve a very large revenue per product family (on the order of $100 million to $1 billion or more), and to have very high “good die” per lot throughput to maximize production capacity. These types of product revenues justify very large design teams, and product development cycles also tend to be short to meet feature upgrade requirements for the market.

In contrast, very high performance networking ICs (such as 100G transceivers) tend to have the opposite development and market criteria. Unit volumes are lower so unit costs are higher and speed-grading is simply not economically feasible due to the low number of wafer runs to support this type of business. The raw silicon is a relatively small part of the final product cost and chip area constraints are not as rigid for these devices. In fact, for thermal performance and to meet the system power budget, it is desirable to increase the die size such that maximum integration and functionality is achieved for the application without using all the available area for DSP logic.

High-end networking ICs are often realized with small, highly specialized teams. They tend to have long development cycles, investment levels that are very high on a per chip basis and life cycles that are much longer (10–15 years). Correspondingly, the investment in design and qualification for networking ICs and systems is high. Upgrades to networks (e.g., from 40 Gbps to 100 Gbps) also tend to be a step-function change in performance and system configuration so that a planned release is very dependent on timely device delivery to ensure a smooth transition. There is usually no fall-back to the previous generation to support some partial functionality.

Taking the above points into consideration, the designer of a 100G IC system and, in turn, a 100G transceiver must carefully consider the potential risks versus rewards of choosing to design in the latest CMOS process technology. They tend to be less stable since models, run sets, design rules and libraries can change. Process variation may be higher or the specified spread may increase later on. Even digital libraries may be immature and may not have optimum power consumption and area. The implications are that over-design is needed to compensate for potential changes in process corner parameters, which usually means increased power and area (digital and analogue), and this may cancel out some of the power saving by using a smaller geometry process.

Longer process turn-around times for new technology nodes, wherein new versions are taped out before previous ones are evaluated, means more design iterations or debug respins are needed. Newer processes always result in higher mask costs initially and design time, cost and effort are all increased compared to the previous process generation. For relatively small markets like 100G long-haul this may not make economic sense.

So it really is necessary to compare all of the pros and cons for such designs. The primary drivers are higher integration (more DSP logic gates) and lower power, but if choosing to design in the latest
process does not result in a significant reduction in power while costing more and increasing time-to-market, then the advantage fades.

6. Conclusions

New ADC and DAC techniques and advances in mixed-signal system integration have now made it feasible to design single-chip CMOS coherent receivers and transceivers which meet the performance and power requirements of long-haul 100G optical systems, and provide a way forward for short-haul and higher-rate applications in future. Realizing these devices is difficult, but the problems can be solved by sufficiently “intelligent design” and do not appear to fundamentally limit their use at higher data rates and in wider application areas.

The biggest challenge in the future is escalating DSP power consumption as data rate and algorithm complexity increases in order to get closer to the Shannon capacity limit of a wider bandwidth channel. This is essentially the same technical problem facing computer CPU designers, but is even more severe due to thermal limitations and the need to fit everything into one chip to minimize the huge volume of inter-chip data traffic which would otherwise be needed.

Without some fundamental advance in digital process and circuit technology and/or algorithms to increase signal processing power efficiency, this is likely to restrict future data rate increases in deployable systems, even if higher rates can be achieved in the laboratory.

References