2 Gbit/s 0.5 μm complementary metal-oxide semiconductor optical transceiver with event-driven dynamic power-on capability

Xingle Wang, Fouad Kiamilev, Ping Gui, Xiaoqing Wang, Jeremy Ekman, Yongrong Zuo, Jason Blankenberg, and Michael Haney

A 2 Gb/s 0.5 μm complementary metal-oxide semiconductor optical transceiver designed for board or backplane level power-efficient interconnections is presented. The transceiver supports optical wake-on-link (OWL), an event-driven dynamic power-on technique. Depending on external events, the transceiver resides in either the active mode or the sleep mode and switches accordingly. The active-to-sleep transition shuts off the normal, gigabit link and turns on dedicated circuits to establish a low-power (~1.8 mW), low data rate (less than 100 Mbits/s) link. In contrast, the normal, gigabit link consumes over 100 mW. Similarly, the sleep-to-active transition shuts off the low-power link and turns on the normal, gigabit link. The low-power link, sharing the same optical channel with the normal, gigabit link, is used to achieve transmitter/receiver pair power-on synchronization and greatly reduces the power consumption of the transceiver. A free-space optical platform was built to evaluate the transceiver performance. The experiment successfully demonstrated the event-driven dynamic power-on operation. To our knowledge, this is the first time a dynamic power-on scheme has been implemented for optical interconnects. The areas of the circuits that implement the low-power link are approximately one-tenth of the areas of the gigabit link circuits. © 2006 Optical Society of America

1. Introduction

Vertical-cavity surface-emitting laser (VCSEL) based optical interconnects for very short reach (VSR) applications have been under active research. High bandwidth, high interconnect density, and low cross talk are among the benefits of VCSEL-based optical interconnects. With the advance of dense optical interconnects based on large VCSEL arrays and photodetector arrays, the power consumption of optical interconnects is gaining increased attention from designers.

Perhaps the most important reason to lower the power consumption of optical interconnects is to alleviate the increasing heat dissipation associated with denser optoelectronic (OE) device packaging. VCSEL devices operating at high temperature display increased jitter, decreased output power, and reduced lifetime. Also, a temperature gradient on large OE device arrays generates performance nonuniformities and complicates compensation requirements. This is because OE device characteristics such as serial resistance, threshold current, and slope efficiency of VCSELs are usually strong functions of temperature. Second, lowering the power consumption of optical interconnects contributes to the overall reduction of the system power budget. As an example, designers of Intel Pentium M microprocessors used a combination of protocol and circuit methods to optimize power consumption of the chip inputs–outputs (IOs).

Power-efficient operation can be explored in various aspects of the system and spans different levels of implementation. Hardware–software trade-offs exist. Static power management, which occurs during the design stage of the system, varies from applying advanced components and architectures in hardware to fixed sleep routines in operating systems. Dynamic power management, on the other hand, tries to min-
optimize the power consumption by adjusting system run-time parameters. For example, dynamic voltage and frequency step-down are two approaches in the power-efficient operations of modern mobile processors. For transceivers, the transmit power is an important parameter that can be dynamically adjusted according to the channel conditions as long as a target performance is maintained. Dynamic transmission power control (TPC) has been extensively studied in wireless channels, and has been done recently in VSR optical interconnects.

In this paper, we consider another significant opportunity for lowering the power consumption of optical interconnects, namely, dynamic power-down and power-on of optical transceivers. The concept is to manage the transceiver operation time by powering it down and waking it up according to user events. These events can be external, such as a communication task or a run-time event managed by the system. While powering down unused blocks is a common practice in power-aware systems, waking up the receiver poses a challenge: there is no communication path that can pass the wake-up event from the transmitter side to the receiver when the transmitter–receiver pair is completely powered down. Continuing to use wireless data transmission as an example, two approaches have been developed in wireless systems to provide the solution. The first is represented by the power-saving mechanism (PSM) in IEEE802.11 standards and many of its improved successors, in which the client (receiver) wakes up periodically to poll for buffered packets at the access point (AP, transmitter). This category is characterized by the extensive overhead of protocols and buffers. The second approach is characterized by an out-of-band, low-power channel that is maintained during transceiver idle times. This channel is used to wake the client up when necessary. The use of a wake-up channel enables the client to move into a deeper powered-down state during idle times and hence achieves a better power-saving performance. The cost is extra radio spectrum.

The concept of a low-power wake-up channel can be readily implemented in optical transceivers to realize a low-cost, event-driven dynamic power-on design. One such implementation is presented in this paper, and we call our technique optical wake-on-link (OWL). Besides the elimination of heavy protocols and hardware overhead associated with IEEE802.11 PSM-like solutions, OWL utilizes the same optical channel at all times in point-to-point links. This is an apparent advantage over similar approaches in the wireless domain, where extra spectrum for the low-power channel is always needed. The rest of the paper is organized as follows: Section 2 discusses the operation and performance of OWL, Section 3 describes our transceiver design that supports OWL, Section 4 describes the system design and implementation, Section 5 presents the experimental results, and Section 6 is the conclusion.

2. General Description of Optical Wake-on-Link

A. Operation

As a low-cost, light-weight, link-level power management scheme, OWL relies on the low-power link maintained during the power-down period of the transceiver, opening opportunities for higher-level tasks to wake up the receiver only when necessary, i.e., only when there is an immediate communication event. OWL is an event-driven scheme. Complex protocols that use prediction to synchronize transmitter/receiver power-on and power-down events are not necessary. A shared optical channel between the normal link and the low-power link greatly reduces the hardware cost because either a fiber channel or a free-space optics channel is an expensive resource. The conceptual operation of OWL is sketched in Fig. 1.

We identify the sequence of events in Fig. 1 as follows. First, both the transmitter (Tx) and the receiver (Rx) sleep in the low-power mode, maintaining a low-power channel. Second, a communication task enters Tx. Third, Tx sends a wake-up message via the low-power channel to Rx. Fourth, Tx turns on into the normal, data communication mode. Fifth, after receiving the wake-up message, Rx turns on into the normal, data communication mode. Sixth, the Tx/Rx pair performs the data communication task. Seventh, on completion of the communication task, Tx/Rx pair turns off into low-power mode, waiting for the next communication task to arrive. It is emphasized that the function that the Rx is remotely awakened by the communication task from the Tx side. While wake-on-link requires a pre-existing low-power channel, turning off Rx is much easier because the power-down command can be embedded in the tail of the communication data.

To this end, we define two modes for the transceiver: active mode and sleep mode. Depending on the external event, the transceiver resides in one of the two modes and switches accordingly. Figure 2 shows the mode transition diagram of OWL where bold arrows indicate sleep-to-active transitions and dashed arrows indicate active-to-sleep transitions. In both transitions, the event enters from the transmitter side and is passed to the receiver by either the normal link or the weak link. The novelty of our transceiver design resides in the fact that it supports a weak link for the receiver remote power-on purpose.

Besides the shared optical channel feature between the normal link and the weak link, the transceiver circuit modules that implement the weak link function were chosen to be physically separated from the main circuit modules that drive the normal link. This is due to the different optimization targets between the normal link and the weak link circuits and is detailed in Section 3.

B. Performance

The performance of OWL is measured by two factors: the power saving efficiency and the data delay overhead. It can be shown that both factors are related to
transceiver transition times between the active mode and the sleep mode. To this end, we define \( \tau_{on} \) as the transition time from sleep mode to the active mode and \( \tau_{off} \) as the transition time from the active mode to the sleep mode. A major challenge that faces a dynamic power-on and power-down scheme in a communication link is the resynchronization cost represented by \( \tau_{on} \). There are several components of \( \tau_{on} \) that are identified as follows: (1) transceiver access time, (2) transceiver DC operation point setup time, (3) amplitude recovery time at the receiver, and (4) clock recovery time at the receiver.

Transceiver access time is measured from the issue of the user “power-on” command (event) to the instant when the “power-on” signal is applied to the controlling element of the transceiver. It is the time for the user to gain access to the transceiver. Transceiver access time involves latencies of system-level signaling and computation and hence is usually the dominant term in \( \tau_{on} \). To reduce transceiver access time, care must be taken not only to the transceiver circuitry, but also to the system design and digital/analog interface of the transceiver. It is possible to use a dedicated IO and data path for passing the “power-on” command. Critical codes like “power-on” can also be stored in read-only memory (ROM) like memory on the transceiver chip. The methods of dedicating special resources to passing “power-on,” however, will greatly increase the hardware cost. We would like to consider the most general case where no more care is taken to passing the “power-on” command than to any other user command or event. In this work, a system command can only be serially passed to the transceiver, and a digital design technique called register preloading is introduced to hide transceiver access time in \( \tau_{on} \). Register preloading is described in Section 4.

Transceiver DC operation point setup time is the turn-on transient time of the circuits. In modern complementary metal-oxide semiconductor (CMOS) technology this time is on the order of nanoseconds. Prebiasing some critical circuits of the transceiver can improve this time, albeit at power consumption penalty.

Primarily driven by research in optical packet switching, amplitude and clock recovery time at the receiver can now be limited to a few nanoseconds as well. Differential signaling shows special advantage here because the signal threshold is embedded in the differential signal, completely removing amplitude recovery time associated with single-ended signaling. For clock recovery at the receiver, gated oscillator-based circuits can be used to implement instantaneous clock recovery. Clock recovery time can also be removed if the transmitter sends a reference clock together with the data. This so-called source synchronization is suitable for VSR applications because clock skew and jitter may be well controlled over short links. It should be noted that the delay overhead of OWL is not more than other dynamic power-on approaches. This is because the transmitter and receiver power-on processes are driven by the same event. PSM of IEEE802.11, on the other hand, does not have this power-on process synchronization mechanism; buffers have to reside at the access point to store the packets that cannot be received by the mobile client instantaneously.

The power saving efficiency of OWL is an additional energy gain built upon the static methods of power-efficient design of the transceiver. A good metric is the ratio of \( E_{saved} / E \), where \( E \) the energy consumed by the transceiver in active mode at all times, and \( E_{saved} \) is the saved energy of OWL with respect to \( E \). The statistics of \( E \) and \( E_{saved} \) are application dependent. For a general analysis they can be estimated by modeling the arrival of events as a Poisson process. In Fig. 3, the event stream is characterized by two parameters: the event duration time \( t_d \) and the interevent time \( t_i \) (idle time). \( t_d \) and \( t_i \) have probability distribution functions as \[ p(t_d) = \frac{1}{\tau_d} e^{-t_d/\tau_d}, \]

\[ p(t_i) = \frac{1}{\tau_i} e^{-t_i/\tau_i}, \]

where \( \tau_d \) and \( \tau_i \) denote mean values of \( t_d \) and \( t_i \), respectively. Assuming that for the worst case, the transceiver consumes the same power during transition times \( \tau_{on} \) and \( \tau_{off} \) as in the active mode, power saving is only achieved during \( t_i - (\tau_{on} + \tau_{off}) \) of one interarrival period \( t_d + t_i \). The average amount of energy saved during one interarrival period is
\[
E_{\text{saved}} = \int_{t_1}^{t_2} \left[ t_i - (\tau_{\text{on}} + \tau_{\text{off}}) \right] \left( P_{\text{active}} - P_{\text{sleep}} \right) \frac{1}{\tau_{\text{on}} + \tau_{\text{off}}} e^{-t_i/\tau_{\text{on}}} \, dt_i = \tau_{\text{on}} e^{-(\tau_{\text{on}} + \tau_{\text{off}})/\tau_{\text{on}}} (P_{\text{active}} - P_{\text{sleep}}),
\]

where \( P_{\text{active}} \) and \( P_{\text{sleep}} \) denote power consumptions in active mode and sleep mode, respectively. The otherwise “always-on” transceiver would consume an average energy during one interarrival period as

\[
E = (\tau_d + \tau_i) P_{\text{active}}.
\]

Plotted against parameters \( \tau_i/(\tau_{\text{on}} + \tau_{\text{off}}) \) and \( \tau_d/(\tau_{\text{on}} + \tau_{\text{off}}) \), Fig. 4 and Fig. 5 show the power saving percentage in three-dimensional surface and two-dimensional curves, respectively. \( P_{\text{sleep}}/P_{\text{active}} \) ratio is set as 1.8/105, roughly the value measured in our transceiver (see below for the experiment data). Parameters \( \tau_i/(\tau_{\text{on}} + \tau_{\text{off}}) \) and \( \tau_d/(\tau_{\text{on}} + \tau_{\text{off}}) \) describe the magnitudes of both the burstiness and busyness of the communication patterns. Note that it is the ratios, not the absolute interevent time and event duration time, that decide the power saving performance, which increases from the lower-right corner to the upper-left corner in Fig. 5. The above results are the worst-case calculations. In cases where the transceiver consumes less power during \( \tau_{\text{on}} \) and \( \tau_{\text{off}} \) there would be better power saving performance.

3. Transceiver Design

For the transceiver, the design objective is to implement a gigabit module for the active mode to perform the normal data communication, and a megabit module for the sleep mode to maintain a weak, low-power link. Current-mode logic (CML) is used for the transceiver gigabit module to achieve high bandwidth and minimum switching noise. While the power consumption of CML circuits does not scale down with the data rate, the requirements for the megabit module are low power and small die area. Hence, the megabit module uses as simple circuits as possible and mostly CMOS logic. The megabit module nevertheless has interface circuits with a board-level signaling scheme, low voltage differential signaling (LVDS) in this case. Furthermore, the gigabit transceiver supports both differential and single-ended optical signaling and is detailed in Subsection 3.C. A conceptual block diagram of the transceiver that supports OWL is shown in Fig. 6.

From this point on, we do not differentiate “gigabit mode” from “active mode,” and we do not differentiate “megabit module” from “megabit transceiver.”

A. Differential Current-Mode Logic Signaling for Gigabit Optical Transceivers

In the signal paths of gigabit Tx and Rx, CML design is used to generate static current consumption and to combat switching noise.\(^{18}\) Block diagrams of the gigabit Tx and Rx are shown in Fig. 7 and Fig. 8, respectively. Some aiding circuits and most of the digital control signal ports are omitted from the block diagrams. Blocks embodied in dashed lines in the receiver can be bypassed or partially bypassed.

The gigabit transceiver accepts and outputs LVDS-compatible electrical signals. At the transmitter, the system LVDS input is amplified by two amplifier (AMP) stages, multiplexed with a differential custom input port, buffered, and then sent to the VCSEL driver. The VCSEL driver is a current-steering differential driver. The biasing and modulation currents are implemented using two binary-weighted current digital-to-analog converters (DACs). Both DACs have 6-bit precision and can produce a current up to 10.16 mA. The receiver consists of two transimpedance amplifiers (TIAs), one AMP filter, two AMPS, one optional delay line stage, one optional retiming...
stage, one buffer stage, and an LVDS driver. The TIA has three digital ports to control and trade-off between the gain and bandwidth according to the received optical power.

B. Megabit Transceiver Design

The objective of the megabit transceiver is to implement a link capable of power code communication at tens of megabits per second. The data rate is sacrificed for low power. Figure 9 is the end-to-end low-power link block diagram. Besides LVDS, the low-power transceivers also support CMOS-in and CMOS-out. If the digital logics embodying OWL scheme are situated near the transceiver, which is the case for our system, CMOS signaling is more convenient than LVDS. We now break down the block diagram and explain the details of low-power operation of the transceivers.

The low-power link is a single-ended link. It only uses half of the optical devices of the differential links used in gigabit mode. The other half is completely turned off during idle times to save power. At the receiver, a static reference current is provided to one input port of the low-power TIA for the recovery of received logic values. The reference current is implemented by a 4-bit current DAC providing a range of 0–128 $\mu$A. The static nature of this current reference makes it necessary to calibrate the link in order to find the optimal reference value.

At the transmitter side, direct or unbiased VCSEL modulation replaces biased VCSEL modulation in gigabit transceivers. It has been well known that biasing a VCSEL above threshold eliminates the initial VCSEL turn-on delay experienced by direct VCSEL modulation.19 The trade-off is made in favor of power consumption here by choosing direct VCSEL modulation. To do this, a CMOS signal is used to turn on/off the VCSEL biasing DAC depending on the current transmitting logic values. The CMOS signal can be either an input or developed from a LVDS input. The low-power transmitter also eliminates two stages of AMPs and one stage of buffer used in the differential gigabit CML transmitter, greatly reducing the power consumption.

At the receiver side, a couple of methods are taken to implement a low-power receiver. First, a passive TIA in Fig. 10 is used to replace the active TIA in the gigabit receiver. This simple passive TIA has only three resistors and three switches controlled by digital bits. The output voltage is developed by PIN photodiode current flowing into one of the three resistors. The large PIN parasitic capacitance and the TIA resistors compose large $RC$ time constants at the input/output nodes, making this method of current-to-voltage conversion only suitable for small bandwidth and hence low data rate requirements.

Second, an underterminated LVDS driver is used to generate the LVDS output. It uses 400-ohm output impedance in contrast to 50-ohm used in the LVDS driver of the gigabit receiver. To understand the trade-off between power consumption and signal integrity of the two types of termination, Fig. 11 shows the diagrams of the two LVDS drivers. The transmission line is 50 ohms—terminated at the far end for both cases. It is apparent that the LVDS driver in the gigabit receiver is able to sink backtraveling waves due to imperfections on the transmission line because of the 50-ohm source termination, thus maximally eliminating intersymbol interference (ISI) due to line imperfections. The LVDS driver is nevertheless only able to inject half of the biasing current into the line, which is the convention of CML design. The 400-ohm unterminated LVDS driver, on the other hand, will incur reflections on the driver’s end, being less immune to ISI due to line imperfections. However, the unterminated LVDS driver is also able to inject a larger portion of the biasing current into the line. In Fig. 11, the gigabit LVDS and the unterminated LVDS are biased at $I_{50}$ and $I_{400}$, respectively. Signaling currents for the gigabit and unterminated LVDS are $I_{50}/2$ and $8I_{100}/9$, respectively. To achieve the same signaling current level, the biasing current of the unterminated LVDS driver is calculated as

$$I_{bias} = \frac{8I_{100}}{9} = \frac{8 \times 100}{9} \mu A$$

$$I_{bias} = \frac{800}{9} \mu A$$

$$I_{bias} = 88.88 \mu A$$
Hence the underterminated LVDS can be biased at a lower level and consumes less power. The output impedance of the underterminated LVDS can be made even larger, virtually infinite, to achieve 100% current injection into the line (standard LVDS driver implementation). The 400 ohms is chosen here based on our experience on ISI requirements.

Third, because of the use of the high-impedance LVDS driver, the three AMP stages in the gigabit receiver are not necessary for generating the LVDS output in the low-power receiver modules.

The LVDS driver is the biggest power consumer in the receiver because it needs to drive system-level signaling lines (on-board traces, for example). In contrast, the CMOS driver is designed to drive nearby digital logic only, thus consuming much less power compared with the LVDS driver.

C. Design of the Embedded Single-Ended Optical Link

Choice of differential or single-ended optical links is a major decision at the beginning of the design process. Each of them has certain advantages over the other. Fully differential optical signaling as shown in Fig. 7, and Fig. 8 is the primary signaling scheme for the gigabit link. However, we would want to test OWL on single-ended links as well. For this reason, we embedded a single-ended link in the differential link. As in Fig. 7, a stage called AMPFLT replaces the first AMP in the receiver. AMPFLT is a fully differential amplifier with the additional capability of processing single-ended input. One terminal of the differential input port of AMPFLT accepts the single-ended input signal, while the other terminal (complimentary terminal) accepts a reference developed by passing the single-ended input through an $RC$ filter. The AMPFLT is designed such that either of the two TIA outputs can be used to develop a signal/reference pair. Figure 12 is the schematics of the AMPFLT.

D. Overview of the Transceiver Chip

An optical transceiver chip was fabricated in the Peregrine 0.5 $\mu$m CMOS silicon-on-sapphire process (SoS). The chip contains 4 transmitters and 4 receivers, supporting 4 point-to-point links when used with another transceiver chip. Each individual link can be in either the gigabit mode or the low-power, megabit mode as in Fig. 6, so that OWL can be tested on different channels to verify functional repeatability. While each link can wake itself up from the low-power mode, a link in low-power mode can also be used to wake up other links into gigabit mode. The gigabit link supports differential and single-ended operations. Optional delay lines and CML retiming flip-flops were inserted in each channel. The die area is 3 mm by 4 mm. Two rectangular regions, each roughly occupying an area of 3.3 mm by 0.5 mm, were allocated to the transmitter and the receiver, respectively. Areas for the low-power modules are about 10% of that of the gigabit modules. The rest of the chip is allocated to pads and bonding areas for OE devices.

The transceiver does not contain digital logic. All digital logic are to be implemented externally to minimize the associated switching noise that may degrade the transceiver performance. Instead, the transceiver provides a register interface to external digital functions. There are 148 registers in total on the transceiver chip, of which 76 belong to the receiver and 72 are transmitter registers. Those registers control the operating conditions of the transceiver by latching users' commands. To minimize the IO numbers of the transceiver, receiver registers and transmitter registers were configured into two serial scan chains.
4. System Design and Implementation

A. Design Partition

Figure 13 shows the functional block diagram of the experiment system. A system command can only be serially scanned into the transceiver through the Din port, and, on completion of the scan, the system issues a level-sensitive Data_Latch pulse to enable the command. Dout port is used for the testing purpose. All digital functions, including OWL demo protocol, were implemented using a Xilinx Virtex II Pro field-programmable gate array (FPGA).21 Besides in situ programmability, the Xilinx FPGA provides Rocket IO high-speed transceiver technologies, which were used to generate high-quality test patterns for the transceiver inputs.22,23 Alternatively, the test signals can be provided by an external data source. On the receiver side, a bit-error rate tester (BERT) was implemented in the FPGA to monitor the link quality. Similarly the receiver output can also be connected to an external data analyzer.

B. System Integration

The physical implementation of the system tried to emulate a backplane environment. As shown in Fig. 14, the system includes a backplane (main board), two line cards (carrier board), two line connectors, and two packages for the transceivers (chip-on-board in our case). Two commercially available 1 × 4 VCSEL arrays and two 1 × 4 GaAs PIN photodetector arrays were flip-chip bonded to the transceiver chips. The VCSEL arrays operate at 850 nm with the typical threshold current at 1.1 mA and the typical slope efficiency at 0.45 mW/mA. The PIN photodetector arrays work with a typical responsivity of 0.5 A/W at 850 nm. The maximum working data rate is 3.6 Gbits/s for the VCSEL and 3.12 Gbits/s for the PIN photodetector, respectively. The device pitch on both arrays is 250 μm.24 Figure 15 is the microphotographs of the transceiver with flip-chip bonded OE arrays. The transceiver integrated circuits with OE devices attached were then wire-bonded to the carrier boards, forming a chip-on-board (COB) structure as shown in Fig. 16. Finally, the two carrier boards were mounted on the main board through two high-speed surface mount connectors. The distance between two carrier boards is 89 mm.

C. Optical and Monitoring System

Figure 17 shows the optical and monitoring system. The system uses free-space optical links. Two high-resolution, seven-element f/2.8 lenses from Universe Kogaku were situated between the transmitter and the receiver, implementing point-to-point links. The first lens collimates the VCSEL beams, while the second lens refocuses them onto the detector array. Between the two lenses, a transparent reflector redirects part of the reflections from the receiver into a CCD camera for monitoring purposes. Hence, the camera views an image of VCSEL lighting spots on the receiver as in Fig. 18. Backlighting in Fig. 17 helps to outline the receiver circuitry onto the camera for better images.

5. Experiment Results

A. Gigabit Link Performance

1. Differential Link Performance

The differential link performance was tested with regards to link BER at 2–3 Gbits/s, performance uniformity spanning different channels, and link immunity to switching noise from adjacent channels. The results are very similar to what we previously demonstrated.18 To summarize, link BER was smaller...

Fig. 14. Schematics of the experiment system.

Fig. 15. Microphotographs of the transceiver with flip-chip bonded OE Arrays: Tx (left) and Rx (right).

Fig. 16. Microphotographs of the chip-on-board structures: Tx (left) and Rx (right).

Fig. 17. Optical and monitoring system.
than $10^{-14}$ at 2 Gbits/s after 11 hours of continuous operation; performance is very uniform across different channels with some channels able to achieve BER < $10^{-14}$ at 2.5 Gbits/s; BER degradation was not observed when imposing worst-case power-on and power-down switching noises from all other channels. We should make a note that the link is a mixed-signal signal path including the transceiver circuits, chip packaging, FPGA fabric, FR4-based printed circuit board environment, and free-space optics. Figure 19 shows a typical eye diagram after 11-hour continuous operation captured at the receiver output. In this case the VCSEL was biased at 1.6 mA. The modulation current was 1.28 mA, generating a peak current of 2.88 mA flowing into the VCSEL. The input is a $2^{31}-1$ nonreturn-to-zero pseudorandom pulse pattern generated by Xilinx Rocket IO.

2. Single-Ended Link Performance
The single-ended link is obtained by enabling the AMPFLT module in Fig. 8. If all other operational conditions are not changed, the single-ended link would process an input signal, the signal-to-noise ratio of which is lower than that of the differential link. Hence we could expect the single-ended link to display a poorer performance in terms of BER. To visualize this, we contrast their eye diagrams in Fig. 20. It can be seen that at 1.5 Gbits/s the difference is not easily visible; at 2 Gbits/s, however, noise starts to smear the single-ended eyes in a noticeable way. In this experiment, VCSEL biasing and modulation current are 1.6 mA and 1.28 mA, respectively.

B. Low-Power Link Performance
The low-power link performance was tested with respect to both LVDS and CMOS signaling. Figure 21 shows the typical eye diagrams at 75 Mbits/s. The low-power link is not used to transmit and receiver data above 100 Mbits/s. In this experiment, the VCSEL was zero biased. The modulation current is 1.2 mA.

C. Gigabit Link and Low-Power Link Power Consumptions
The gigabit link and low-power link display different power consumption patterns. While the gigabit link draws static current from power supplies at all times, the low-power link power consumption is both data-rate dependent and data-pattern dependent. If the data is DC balanced, the data-rate dependency of the low-power link power consumption is minimized because the VCSEL’s on and off times are roughly the same for all data rates. Figure 22 shows the measured power consumptions for gigabit and low-power links (transmitter and receiver combined, single channel) as described in Subsections 5.A and 5.B.

OWL requires that the transceiver be in low-power mode during idle times. Low-power transceiver biasing currents determine the power consumption floor of the transceiver. The power consumption floor is 3.8 mW for LVDS and 1.8 mW for CMOS, about two orders of magnitude lower than that of the gigabit link.

D. Optical Wake-on-Link Experiment

1. General Description
The objective of this experiment is to demonstrate receiver wake-on-link capability. In Fig. 23 an OWL controller is implemented in the FPGA. The controller generates and sends three types of events (commands) to the transceiver. First, link power-down is issued electrically to both the transmitter and the receiver. After receiving this command, the transceivers switch to the sleep mode, establishing the megabit, low-power link. Second, Rx wake-up is is-
sued to the transmitter and sent to the receiver via the low-power link. After receiving this command, the receiver powers on to the gigabit mode. Third, Tx wake-up is issued to the transmitter. After receiving this command, the transmitter powers into gigabit mode. The gigabit link is then established. The formats and timings of the three commands are all controllable.

The experiment uses CMOS signaling of the low-power link. Bits traveling on the low-power link are driven by a 50 MHz clock. For visualization purpose the commands are implemented as a “111” sequence that displays as a 60 nS positive pulse on the oscilloscope. For clarity purpose, the demo protocol does not acknowledge the wake-up message on the receiver side.

During the experiment, 2 Gbit/s pseudorandom data were constantly fed into the transmitter. Depending on the transceiver mode, there may or may not be output data from the receiver. This provides us with information on whether the desired mode transition has occurred. Receiver wake-on-link function was tested against both differential and single-ended optical links of the gigabit mode. Wake-on-link capability was also tested across different channels. For example, the receiver power-on can be induced by a wake-up message conveyed by another channel. All the tests were successful.

2. Register Preloading to Hide Transceiver Access Time in $\tau_{on}$

As previously stated in Section 2, transceiver access time is usually the dominant portion in $\tau_{on}$ because the “power-on” command is of digital bits and must be propagated and transformed to a simple controlling signal applied to the transceiver. Figure 13 shows that passing a system command to the transceiver is a two-phase procedure: (1) reconfiguration of all 72 transmitter registers or 76 receiver registers, and (2) issue of Data_Latch pulse. Phase (1) in turn dominates the transceiver access time because the registers are configured serially. Denoting the low-power link system clock period as $T$, phase (1) then requires $76T$ ($72T$ for the transmitter) to complete and phase (2) requires $1T$ to conclude. One way to get around this problem is to move phase one of both Rx wake-up and Tx wake-up up to immediately after the completion of the link power-down command. Rx wake-up and Tx wake-up commands can then be reduced to only phase (2). This technique, which we call register pre-loading, greatly reduces the transceiver access time in $\tau_{on}$. This technique can be used to any register serial scan chain as long as the next configuration of the scan chain is known.

3. Demo of Optical Wake-on-Link

Figures 24 and 25 are the typical results of two slightly different experiment schemes. The system command stream is displayed in the uppermost waveform. The waveforms in the middle and bottom show the outputs of the receiver. The register pre-loading technique moves all register chain scan operations to after the link power-down command. Hence no delays incurred by the digital logic are present on command 2 (Rx wake-up) and 3 (Tx wake-up). In Fig. 24, the link power-down command was issued first to power down the link into low-power mode. Rx wake-up command was then issued. The validity of Rx wake-up command was not known until command 3, Tx wake-up, after which the receiver was able to capture the 2 Gbit/s pseudorandom data, indicating that the receiver had indeed been wakened up into the gigabit mode. The interval between command 2 and 3 is not necessary. In Fig. 25 they are merged into one command, and the link can still be awakened into the gigabit mode.

4. Power Saving Efficiency Calculations

This subsection gives the calculation of power saving efficiency based on patterns in Fig. 25. Event duration time $\tau_d$ and interevent time $\tau_i$ in this experiment were chosen to be $78T$ and $200T$, respectively, where
There are several interesting directions for the future work. First, due to physical size limitation, the VCSEL array and the photodetector array have not been flip-chip bonded onto the same transceiver die. This work has been based on unidirectional optical links. On the other hand, bidirectional communication may open opportunities for more elaborate dynamic power management methods due to the hand-shaking capability between the two communication peers. One possibility is adding transmission power control to minimize the power consumptions of the transceiver active mode and sleep mode. Second, OWL is an event based power-down and power-on technique. It is possible to extend OWL to packet based optical transceivers to realize a complete data driven power-down and -on technique. The challenge will be for the sleep mode to provide fast response times when a packet arrives as well as to maintain low power consumptions during interpacket times.

The authors acknowledge the support from DARPA through the C2OI program under the W911NF-04-1-0035 subcontract. We thank Michael J. McFadden at the University of Delaware for providing help on the optical setup in the experiment.

References

11. C. F. Chiasserini and R. Rao, “Combining paging with dynamic

$T$ is the clock cycle time (in this case $T = 20 \text{ ns}$, corresponding to a 50 MHz clock). From the discussion in Subsection 5.C.2, we identify both of $T_{on}$ and $T_{off}$ to be $77T$. The register preloading technique moves $76T$ of $T_{on}$ to immediately after $T_{off}$ leaving effective $T_{on}$ to be only $17T$ (DataLatch pulse time). Note that Fig. 25 simply labels $T_{on}$ after $T_{off}$. In this system, power consumption during $T_{on}$ is the same as in $T_{i}$ which defers from the worst-case discussed in Subsection 2.2. We define

\[
\begin{align*}
    x &= \frac{T_{d}}{T_{off}} = \frac{78T}{77T} = 1.013, \\
    y &= \frac{T_{i}}{T_{off}} = \frac{200T}{77T} = 2.597.
\end{align*}
\]

Based on Section 2 we then calculate the power saving efficiency as

\[
\frac{E_{saved}}{E} = \frac{y}{x+y} \exp\left(-\frac{1}{y}\right)\left(1 - \frac{1.8}{105}\right) = 48.1\%.
\]

Therefore OWL applied to patterns in Fig. 25 saves nearly half of the power that otherwise would be dissipated by an “always-on” transceiver. Intuitively, this result can be understood by noting that in Fig. 25 the gigabit PRB data section and the sleep-mode section (the “straight line”) are approximately of the same length.

6. Conclusion

A low-cost, light-weight, event-driven transceiver power-on scheme was proposed and implemented for optical interconnects. This scheme, optical wake-on link (OWL), utilizes a megabit, low-power link maintained during idle times to wake up the receiver on the arrival of a communication event. During idle times, the low-power mode provides a power-consumption floor of 1.8 mW on a channel using CMOS signaling, while the channel’s normal, gigabit mode typically consumes over 100 mW. A 0.5 μm CMOS parallel optical transceiver chip and a backplane-based system were designed and tested. The experiment successfully demonstrated OWL operation.

![Fig. 25. Optical wake-on-link demo: Scheme 2.](image)
12. E. Shih, P. Bahl, and M. Sinclair, “Wake on wireless: an event-
driven energy saving strategy for battery-operated devices,” in
ACM/IEEE Proceedings of the Eighth Annual International
Conference on Mobile Computing and Networking (IEEE,
power-on for VCSEL-based gigabit optical transceivers,” in
314–315.
receiver and its applications in optical multiaccess networks,”
15. Y. Ota, R. Swartz, V. Archer, S. Korotky, M. Banu, and A.
Dunlop, “High-speed, burst-mode, packet-capable optical re-
ceiver and instantaneous clock recovery for optical bus opera-
16. M. Banu and A. E. Dunlop, “Clock recovery circuits with
instantaneous locking,” Electron. Lett. 28, 2127–2130
17. D. Bertsekas and R. Gallager, “Delay models in data net-
works,” in Data Networks (Prentice-Hall, 1992), pp. 149–
270.
18. P. Gui, F. Kiamilev, X. Wang, X. Wang, M. McFadden, M.
Haney, and C. Kuznia, “A 2Gbps 0.5μm CMOS parallel optical
transceiver with fast power-on capability,” J. Lightwave Tech-
19. R. G. Hunsperger, “Direct modulation of semiconductor las-
20. Peregrine Semiconductor SoS CMOS Foundry Service; http://
Document UG012; http://www.xilinx.com/bvdocs/userguides/
UG024; http://www.xilinx.com/bvdocs/userguides/ug024.pdf
(24 February 2004).
24. Emcore Corporation, "1 × 4 VCSEL array 2.7–3.6Gb/s, 8685-
1402," and "1 × 4 GaAs PIN photodiode array model 8485-