Optical packet transmission in 42.6 Gbit/s wavelength-division-multiplexed clockwork-routed networks

E. Bravi, A. D. Ellis, A. K. Mishra, and D. Cotter*

Photonic Systems Group, Tyndall National Institute, Lee Maltings, Prospect Row, Cork, Ireland
**Corresponding author: david.cotter@tyndall.ie*

Received October 29, 2007; revised January 23, 2008; accepted January 23, 2008; published March 18, 2008 (Doc. ID 89098)

The use of amplitude-modulated phase-shift-keyed (AM-PSK) optical data transmission is investigated in a sequence of concatenated links in a wavelength-division-multiplexed clockwork-routed network. The narrower channel spacing made possible by using AM-PSK format allows the network to contain a greater number of network nodes. Full differential precoding at the packet source reduces the amount of high-speed electronics required in the network and also offers simplified header recognition and time-to-live mechanisms. © 2008 Optical Society of America

OCIS codes: 060.4250, 060.4510.

1. Introduction

The performance and scalability of the network interconnecting large numbers of processors is one of the key challenges in the future development of the most powerful supercomputers. Optical networks offer the possibility of large bandwidth and low latency, provided stringent cost constraints can be met. Various optical packet-switched network designs have been proposed for this application, including the packet self-routing network called data vortex [1–3], the optical broadcast-and-select network called OSMOSIS [4,5], and toroidal networks using wavelength-division-multiplexed clockwork routing (WDM-CR) [6–8]. It has been shown, by discrete-event traffic simulation, that WDM-CR is capable of achieving the primary design goals of large scalability (to several thousands of optical network nodes) and low latency [7,8]. This paper focuses on the optical transmission aspects of the WDM-CR network. The main observations reported here may be applicable to other densely multiplexed optical network architectures in which signals traverse several concatenated transmission links. In particular we demonstrate experimentally a channel spacing and number of concatenated links that would allow the scalability of an optical packet-switched interconnection network to a level not previously achieved. In addition, we identify and analyze, for the first time to the best of our knowledge, significant error-multiplication effects in concatenated transmission links using duobinary modulation and propose techniques to utilize these effects.

2. Packet Transmission in WDM-CR Networks

WDM-CR is a synchronous optical network design, optimized for transmission of short (~240 bytes) packets with minimal latency. A full description of the architecture has been given previously [6,7], and only a brief summary of certain relevant features is given here. The packets are transmitted at high speed (42.6 Gbits/s) in fixed-duration time slots (~50 ns), and the network is synchronized to a global clock. The network nodes are arranged in m densely interconnected two-dimensional (2D) tori, as illustrated in Fig. 1, and the total number of nodes in the network is n^2m, where the tori are each of dimensions $n \times n$. A total of $2m$ wavelengths are employed in the network, consisting of m pairs: (λ_1,λ_1'), (λ_2,λ_2'), $\ldots$, (λ_m,λ_m'). Each pair corresponds to one of the m tori [for example, (λ_q,λ_q') is used for transmission within the qth torus], and the two wavelengths in a pair correspond to transmission in the row and column directions of

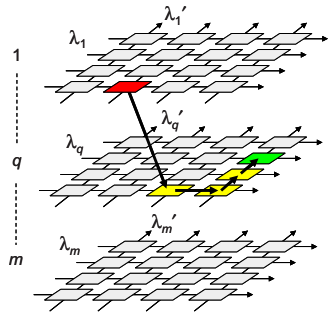

Fig. 1. 3D representation of the WDM-CR network topology [7], comprising m densely interconnected 2D $n \times n$ tori. For simplicity, each 2D torus is illustrated by a planar grid representing a small portion of its surface. Here, as an example, a source node (red) on the first torus transmits a packet to a destination node (green) on the qth torus, utilizing four links via three intermediate nodes (yellow).

the torus. A packet may be transmitted from any node to any other node in the network. The entire route to be taken by a given packet is predetermined at the source by the particular choice of three parameters: the outbound link direction from the source, the packet time slot within a continuously repeating frame, and the outgoing optical wavelength. This wavelength, which is any one of the $2m$ wavelengths used in the network, is selected at the source using a fast tunable laser. In the example shown in Fig. 1, the packet is transmitted on the first link from the source with wavelength λ_q and is thus directed to the qth torus, in which the destination resides. After the first link, the packet is routed automatically toward its destination by an autonomous process called clockwork routing (CR), which has been fully described previously [6,7,9]. In the example shown in Fig. 1, at intermediate nodes the packet is transmitted onward automatically with wavelength λ_q or λ_q', depending on whether the current link is aligned in the row or column direction. This WDM-CR routing scheme ensures that the choice of the three routing parameters at the source is sufficient to determine the entire path of a packet from its source to destination—over several successive transmission links—without further active intervention or rerouting needed by intermediate nodes, apart from determining whether the destination has been reached. Contention within the network is entirely avoided by holding the packet in an electronic buffer at the source until a suitable time slot becomes available (as described in [6,7,9], the choice of time slot ensures correct automatic CR to the destination). The subsequent transmission path is then guaranteed to be contention free, without the need for prior signaling, reservations, buffering at intermediate nodes, or routing deflections.

All packets incoming to a node are converted to the electrical domain, and thus packets in transit undergo optical–electrical–optical (O-E-O) conversion at every node. Therefore, from a transmission perspective, the network consists of concatenated point-to-point links. This has several advantages [7]: The optical design is greatly simplified; engineering and performance issues due to packet-to-packet power and optical signal-to-noise variations are avoided; and additionally, packet-by-packet clock recovery is greatly simplified for single point-to-point links.

Figure 2 is a simplified schematic of the transmission path taken by a packet from its source to destination in a WDM-CR network. (The path shown is only a representative portion of the full 3D optical architecture, which is fully described in [7]. This simplification is made here in order to focus on the optical transmission aspects of the WDM-CR network.) The routing table at the source provides values of the three routing parameters described above, and the packet is inserted into the appropriate electronic buffer to await transmission over k successive links to its destination; here k has a predetermined value known at the source, where $1 \leqslant k \leqslant 2(n-1)$ and the toroid has dimensions $n \times n$. In the simple example shown in Fig. 1, $k=4$. In the case of a 2048-node network constructed on an 8×8 toroid utilizing 32 wavelength pairs [7], the maximum number of links traversed by any packet is 14.

As shown in Fig. 2, the packet is transmitted on the first link via a tunable transmitter (TTX), which is capable of being tuned to any one of the $2m$ wavelengths used in the network. The TTX is tuned to the appropriate wavelength according to the routing table at the source; depending on the selected outgoing link direction, this is one

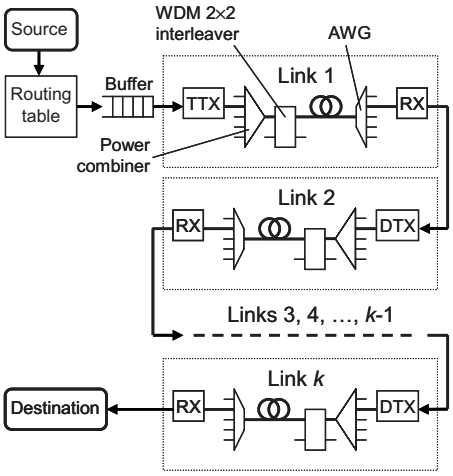

Fig. 2. Schematic of concatenated transmission links in a WDM-CR network.

of a pair, say (λ_q, λ_q'), corresponding to the torus in which the packet destination resides. Each subsequent link uses a switched dual-wavelength transmitter (DTX) which, in a given time slot, transmits on either one of the wavelengths λ_q and λ_q'. In CR the wavelength of every DTX is switched in a fixed periodic fashion in synchronism with a global time-slot clock [6]. Depending on the wavelength of the DTX, the packet is forwarded on one or another of the two outputs of a WDM 2×2 interleaver (corresponding to the row and column directions of the toroid). In the full WDM-CR architecture described in [7], each network node is equipped with at least one TTX in order to function as a packet source, as well as two DTXs to forward packets already in transit. However the path followed by a particular packet, as shown in Fig. 2, consists of a first link using a TTX, and $(k-1)$ subsequent links using DTXs.

Other components encountered in each link in Fig. 2 are a power combiner (which combines the outputs of several TTX and DTX transmitters) and an arrayed waveguide grating (AWG) WDM demultiplexer (which directs the packet to the required receiver RX). In a fully populated WDM-CR network, many packets at different wavelengths may propagate simultaneously in the interleaver and demultiplexer in each link, and therefore adjacent channel cross talk is a limiting impairment. However, a particular design feature of the network is that the channel spacing at the input to the AWG is twice as wide as that at the input to the interleaver, and therefore the possibility of adjacent channel cross talk is greatest at the interleaver. A more detailed explanation of the routing functions of these components and the overall WDM-CR routing scheme has been presented previously [6,8]. The overall scalability of the WDM-CR network (in terms of the number of network nodes) is determined mainly by the number of wavelength channels that can be selected by the tunable transmitters. This number is maximized by using the narrowest possible channel spacing, consistent with error-free optical transmission.

3. AM-PSK Transmission over Multiple Concatenated Links

The experiment described in Section 4 investigates transmission of 42.6 Gbit/s optical packets through a sequence of links such as those shown in Fig. 2, each containing a WDM 2×2 interleaver with 50 GHz channel spacing. Preliminary bit-error-rate (BER) measurements, to be described in Section 4, showed that, when using standard non-return-to-zero (NRZ) intensity modulation, excessive system penalties were observed due to adjacent-channel cross talk at this narrow spacing. Therefore, to reduce these penalties and ensure error-free transmission over multiple links, amplitude-modulated phase-shift-keyed (AM-PSK) coding has been adopted for use in the WDM-CR network. It is well known that AM-PSK modulation has smaller transmission bandwidth compared to binary intensity modulation for the same information throughput [10]; typical examples of measured optical spectra are shown for comparison in Fig. 3. Previously the reduced transmission bandwidths of both three-level optical duobinary signals and AM-PSK signals have been utilized to reduce the chromatic dispersion in long-distance optical transmission [11–13], to increase the information

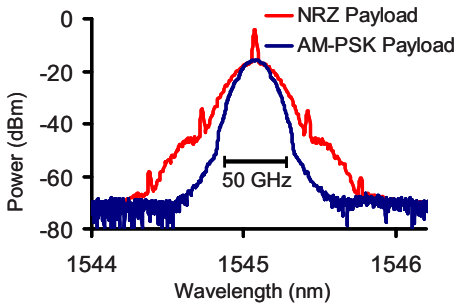

Fig. 3. Typical measured optical spectra of 42.6 Gbit/s signals using NRZ intensity modulation and AM-PSK modulation.

spectral density in high-capacity dense WDM transmission systems [14], or to increase the throughput in a wavelength-routed switch [15]. In this work, the narrower spectrum is utilized to increase the number of wavelength channels available, so that the WDM-CR network can accommodate a greater number of nodes [6,7]. The AM-PSK modulation format was chosen, since in that case the receiver can consist simply of a standard intensity-modulation direct-detection (IM-DD) receiver followed by a binary data inverter [13], as opposed to a three-level receiver [11]. This is because in AM-PSK, the amplitude modulated (AM) part of the modulation carries all of the information and is detected as a binary signal, while the phase-shift-keyed (PSK) part serves only to compress the spectrum and the phase levels are not detected. AM-PSK was selected for this application in preference to other modulation formats with reduced spectral width (such as vestigial sideband, differential quadrature phase-shift keying, orthogonal frequency-division multiplexing, and phase-polarization-diverse schemes) because AM-PSK minimizes the overall complexity and latency of the system, as well as minimizing component count and cost.

There are three approaches that can be used to implement end-to-end AM-PSK transmission in the multilink system shown in Fig. 2. The first approach is to digitally recode the binary data signals at the full transmission line rate at each intermediate node; this approach maximizes the amount of high-speed electronics needed in the nodes. A second approach is to perform full differential precoding of the data packets at the source before insertion into the input buffer, in preparation for AM-PSK transmission over k sequential links, without the need for digital recoding at the intermediate nodes. In this case each intermediate node can simply incorporate a suitable electrical low-pass filter to perform analog conversion of the output from the IM-DD receiver to duobinary format, for onward transmission in AM-PSK format after optical modulation. In the third approach, the signals are transmitted without any coding, and the resulting data scrambling caused by the repetitive AM-PSK modulation and detection is reversed by using appropriate decoding at the final receiver. This third approach is not readily suitable for a WDM-CR network, because the final receiver does not have prior knowledge of the number of modulation stages through which a packet has passed. However, the second approach—performing full differential precoding at the source—is suitable for use in WDM-CR networks because the whole transmission path (and hence the number of links k) for a packet is predetermined at the source, and the precoding operation may be performed prior to entry into the high-speed buffer. This has the advantage of requiring less high-speed electronics in the nodes, and also, as described below, opens the possibilities of simplified header address recognition and time-to-live signaling.

The implementation of multilink AM-PSK transmission with full differential precoding at the source is now described. We first recall the conventional procedure for duobinary transmission over a single link [10]. Before conversion to duobinary format, the binary data is precoded, so that the coded bits are given by $c_j = c_{j-1} \oplus d_j$, where d_j is the jth data bit, and $\oplus$ denotes subtraction modulo 2. These coded bits are then converted to three-level symbols t_j, given by $t_j = c_j + c_{j-1}$, which take values 0, 1, and 2. This conversion, which corresponds to a delay-and-add digital filter, is often implemented using a simple low-pass analog filter. In an optical AM-PSK transmission system, this three-level signal t_j is then further processed using a Mach–Zehnder modulator (MZM), such that $t_j = 1$ corresponds to zero light intensity, and $t_j = 0, 2$ corresponds to transmission of optical electric fields of equal amplitude but opposite

sign [13]. At the receiver, an optical square-law detector converts $t_j=0,2$ signals to binary 1 values and $t_j=1$ to binary 0. Effectively, this overall scheme of data precoding, low-pass filtering and optical AM-PSK modulation at the transmitter, together with square-law detection at the receiver, results in the receiver sampling the value $c_j \oplus c_{j-1}=c_{j-1} \oplus d_j \oplus c_{j-1}=d_j$.

It follows that for AM-PSK transmission over k concatenated links the necessary precoding is given by the recursive expression

$$(c_j)_k = (c_{j-1})_k \oplus (c_{j-1})_{k-1} \oplus \cdots \oplus (c_{j-1})_1 \oplus d_j = (c_{j-1})_k \oplus (c_j)_{k-1} \quad \text{for } k \geq 2, \tag{1}$$

where $(c_j)_k$ denotes the jth binary symbol precoded for transmission over k links. Then, in the absence of any errors, a receiver located after h links ($1 \leq h < k$) samples the signal:

$$(c_j)_{k-h+1} \oplus (c_{j-1})_{k-h+1} = (c_{j-1})_{k-h} \oplus (c_{j-1})_{k-h-1} \oplus \cdots \oplus (c_{j-1})_1 \oplus d_j = (c_j)_{k-h}. \tag{2}$$

At the destination ($h=k$), provided again that the system is error-free, the receiver samples the value $(c_j)_1 \oplus (c_{j-1})_1 = d_j$, thus recovering the original data.

The way in which errors accumulate in the AM-PSK transmission system differs from that in a binary system. In a binary system comprising concatenated transmission links, errors occurring in one link will propagate onward through subsequent links without modification, and therefore the signal reaching the final receiver contains a linear accumulation of the errors occurring in all the links. However, in the case of AM-PSK transmission, a transmitted symbol depends on the values of both the current bit and the previous bit, and this interaction therefore causes any bit errors to multiply in number in a system of concatenated links. For example, if the received signal on any link contains a single isolated bit error (in the pth bit position, say), this will result in two errors being received on the next link (in bit positions p and $p+1$), and these two errors then propagate on through subsequent links producing further errors. After propagation over q links since the first occurrence of the isolated bit error, a total of $q+1$ bits are affected by the initial error (the bits in positions p, $p+1, \ldots, p+q$). However, as shown in Fig. 4, not all of these affected bits may have erroneous values; this is because errors in adjacent bit positions may interact in the next stage of transmission so as to correct an error. For example, on the second link after the first occurrence of a single error, three bits are affected by the original error (the bits in positions p, $p+1$, and $p+2$), but the received signal contains only two errors (errors occur in bit positions p and $p+2$, but not in position $p+1$). Thus a single isolated error gives rise to a fixed deterministic pattern of bit errors in the received signals on subsequent links, and this fixed pattern is shown in Fig. 4. A notable feature of this pattern is that, after q links, the bits in positions p and $p+q$ are always

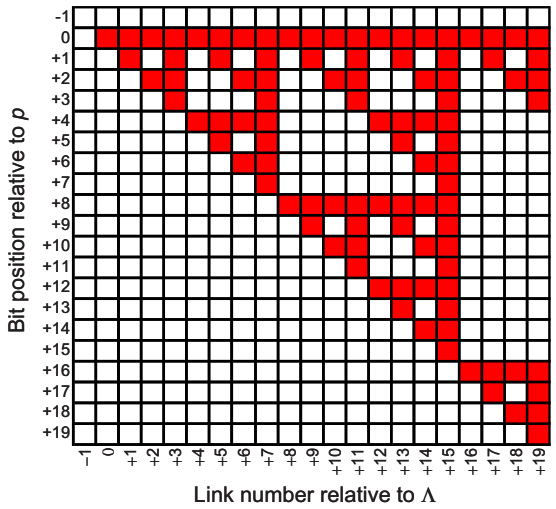

Fig. 4. Diagram showing the evolution of bit errors in a sequence of concatenated links using AM-PSK transmission. A red rectangle indicates a bit error in the receiver signal on a given link. A single isolated error, in bit position p on link Λ, results in a characteristic pattern of errors on subsequent links, as shown. This fixed pattern is independent of the values of the binary data.

errored. As described in Section 6, this special property of concatenated AM-PSK transmission can be exploited as a time-to-live mechanism for optical packets.

4. Transmission Experiment

Experiments were performed to verify the characteristics of optical data packet transmission through a sequence of links, as in Fig. 2. First, however, preliminary BER measurements were performed on continuous or long bursts of 42.6 Gbit/s data in WDM channels spaced by 50 GHz using NRZ intensity and AM-PSK optical modulation formats to determine which of these formats is suitable for transmission through the sequence of links. The results of these preliminary measurements, shown in Fig. 5, indicated that the single-channel back-to-back performance of NRZ and AM-PSK was almost identical (the receiver sensitivities differed by only 0.34 dB), as expected [13]. However, after insertion of a WDM 2×2 interleaver with 50 GHz channel spacing and in the presence of adjacent WDM channels, the receiver sensitivity suffered a 9.7 dB penalty when using NRZ modulation, due mainly to cross talk, whereas in the case of AM-PSK format the penalty was only 1.3 dB, due to the lower cross talk resulting from the narrower signal spectrum. These preliminary measurements verified that the AM-PSK format is more suitable for the WDM-CR network with 42.6 Gbit/s optical data packets and 50 GHz channel spacing. In contrast, NRZ modulation is unsuitable for 50 GHz channel spacing and would require 100 GHz routing optics. Thus, compared to NRZ, AM-PSK allows a twofold increase in the number of wavelength pairs m that can be utilized [15], and since, as noted earlier, the number of nodes in the network is n^2m, this doubles the number of nodes.

The experiments to verify the transmission characteristics of 42.6 Gbit/s optical packets through a sequence of links were performed using the arrangement shown in Fig. 6. The experiments simulated the switching of data packets between the row and column directions of the WDM-CR network toroid, as well as transmission over a variable number of concatenated links. Data packets were generated using a standard NRZ optical transmitter, which was not optimized for packet operation. The NRZ pattern generator was programmed with a 150 bit guard band followed by a 256 byte data packet (and further pseudorandom 256 byte words during unused time slots). In practice, a packet stream was generated using a $2^{13}-1$ pseudorandom bit sequence, which was coded into RS (255,243)(243,231) blocks (10% coding overhead), with the first 150 bits set to zero. This packet stream was transferred to the internal memory of the pattern generator. This optical signal was then detected by the optically preamplified photodiode PD1, which thus represented the receiver on the first (or later) link of the sequence in a WDM-CR network (Fig. 2). The NRZ formatted data at the output of PD1 was amplified and split to provide an input to a clock recovery unit (CRU) and a *D*-type flip-flop (DFF1), which retimed the data. This retimed signal was converted to a three-level signal using a duobinary drive amplifier (DDA1), which then drove a MZM (MZM1). The MZM modulated the output of the first dual-wavelength switched

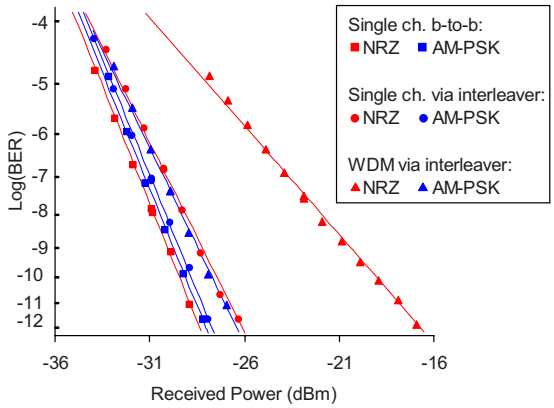

Fig. 5. Measured bit error rates of continuous 42.6 Gbit/s data using NRZ and AM-PSK modulation formats in the following cases: single channel back-to-back (b-to-b), single channel after transmission through a WDM 2×2 channel interleaver (50 GHz channel spacing), and after transmission through the same interleaver in the presence of adjacent WDM channels.

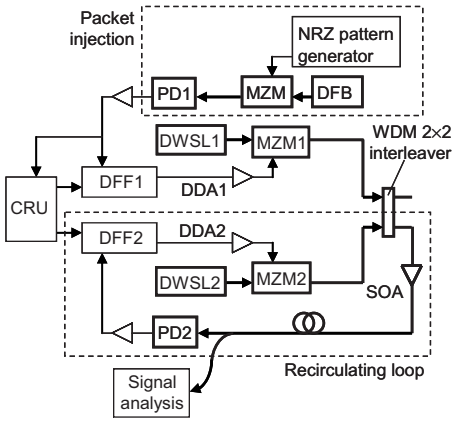

Fig. 6. Experimental arrangement to verify 42.6 Gbit/s optical packet transmission over multiple concatenated AM-PSK links. The recirculating loop emulates transmission over successive identical links, as in Fig. 2. DFB: distributed feedback laser; MZM: Mach–Zehnder modulator; PD: photodiode; DFF: *D*-type flip-flop; DDA: duobinary drive amplifier; DWSL: dual-wavelength switched laser; SOA: semiconductor optical amplifier; CRU: clock recovery unit.

laser (DWSL1) to produce an AM-PSK signal [13]. Each of the DWSLs shown in Fig. 6, nominally centered at 193.3 THz, was tuned such that they could be switched rapidly between two wavelengths, λ and λ', corresponding to "odd" and "even" wavelengths of the WDM 2×250 GHz channel interleaver, to thus direct signals to the required output port (equivalent to row and column directions of a torus in the WDM-CR network). DWSL1 was initially set to the wavelength, which directed the data packet through the interleaver towards PD2, where it was detected [as shown in Fig. 7(a)]. The output of PD2 was directly connected to DFF2 to minimize low frequency distortions and to retime the data. Meanwhile the state (wavelength) of both DWSLs was changed, with timing such that the wavelength switching of DWSL2 coincided with the arrival of the 150 bit guard band at MZM2. The binary data signal from DFF2 was passed through DDA2, so that MZM2 reconverted the signal to optical AM-PSK format. Figure 8(a) illustrates (from left to right) the attenuation of 50 trailing bits of data transmitted from MZM1 caused by the switching of DWSL1, a switch-over transient of DWSL2 within the guard band, and the first part of the packet transmitted from MMZ2 via the interleaver back towards PD2.

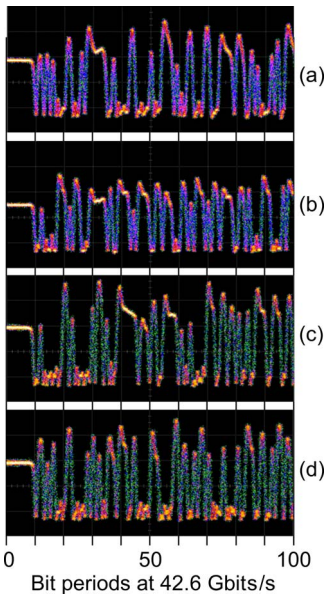

Fig. 7. Leading part of received data packets after transmission through a sequence of links: (a) after first insertion into the recirculating loop; after (b) 1, (c) 8, and (d) 15 circulations. The full horizontal scale width corresponds to 100 bits at 42.6 Gbits/s.

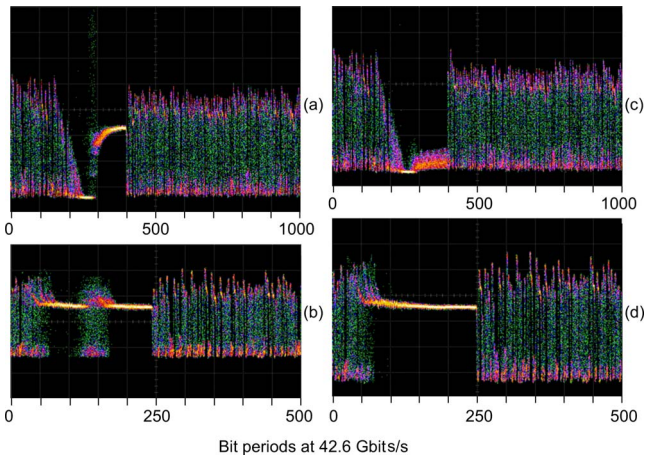

Fig. 8. Leading part of received data packets after transmission through a sequence of links, showing the effects of different preamble patterns: (a),(b) all zeros; (c),(d) 101010. The traces are recorded (a),(c) after first insertion into the recirculating loop and (b),(d) after eight circulations. The full horizontal scale widths correspond at 42.6 Gbits/s to (a),(c) 1000 and (b),(d) 500 bits.

Thus, in these experiments, DWSL1 was used to switch packets into the recirculating loop shown in Fig. 6, while DWSL2 was used to hold a packet in the loop during several circulations. In effect, DWSL1 and DWSL2, in combination with the 2×2 interleaver, performed the switching function for the recirculating loop, which more conventionally is performed by an optical space switch. On each pass around the loop, the packet was received and converted to binary format in PD2, passed through DFF2 for retiming, reconverted to AM-PSK optical format via DDA2 and MZM2, and routed back through the 2×2 50 GHz channel interleaver. BERs and eye patterns were recorded with conventional equipment using an optical signal tapped from the input to PD2. For the purpose of BER measurements on data packets, the length of fiber in the recirculating loop was increased to 400 m to ensure correct operation of the pattern synchronization of the error detector, enabling detailed characterization of the error characteristics of the transmission links. However, in a WDM-CR network, the length of every link is required to match the fixed time-slot duration (typically 50 ns, approximately equivalent to 10 m of fiber, in the case of the supercomputer application [6,7]).

The preliminary BER measurements on continuous data signals, mentioned earlier, had determined that the presence of an AWG (with twice the channel spacing of the interleaver) in each transmission link (Fig. 2) had negligible impact on the performance in the case of AM-PSK signals. Also, using the results of the preliminary BER measurements and the known insertion losses of the WDM components, a calculation of the power budget and optical signal-to-noise ratios indicated that optical amplification would be required in the links of a large-scale WDM-CR network to ensure error-free operation. For these reasons, a semiconductor optical amplifier (SOA) was incorporated as a power amplifier in the transmission experiment of Fig. 6, to ensure inclusion of noise and nonlinear signal impairments, but for simplicity the AWGs and passive power combiners were omitted. In the experiments, the point-to-point regenerated transmission was limited by thermal noise in the receiver, rather than optical noise, and therefore the SOA was used to increase the received signal power above the thermal noise limit. A SOA is preferred over a fiber amplifier for this application, because it is more compact and does not significantly increase latency.

For this experiment, the DWSL was constructed from two distributed feedback (DFB) lasers, one at an odd wavelength of the interleaver and the other at an even wavelength), driven by complementary drive currents such that one laser was held below threshold while the other lased, and these two outputs were combined to give a single switched dual-wavelength output. A small guard band of ~5 ns was introduced between the switch-down of one drive current and switch-up of the other, to minimize optical coherent interference effects during the switch-over. The lasers used were commercial devices designed for direct modulation at 2.5 Gbits/s, and the rise and fall times of the switched laser outputs were measured to be ~2.2 ns, limited by the cur-

rent driver. A potentially lower-cost approach that could be used in the future for the DWSL would be a single semiconductor device, such as a multicontact laser optimized to switch between two wavelengths or a slotted Fabry–Perot device customized for dual-wavelength switching [16].

5. Experimental Results

Figure 7 shows oscilloscope traces of the leading part of the received data packet after injection into the recirculating loop [Fig. 7(a)] and after several circulations [Figs. 7(b)–7(d)]. [We note that the data patterns shown in Figs. 7(b)–7(d) all resulted from the same input data pattern, Fig. 7(a)]. For the known input data patterns, it was verified by inspection that the observed bit patterns received after h circulations (for $h=1,2,\ldots,15$) were accurately predicted by expression (2), after taking account of the fact that the signal outputs from the photodetectors PD1, PD2, and DFF were noninverting. The measured BER versus the number of circulations (Fig. 9) show error-free operation ($<10^{-9}$) for signal powers incident on the photodetector of +8 dBm. Degradation due to saturation of the photodetector is expected for powers higher than $\sim+10$ dBm. As noted earlier, the point-to-point regenerated transmission was limited by thermal noise, and therefore there is no accumulation of noise from link to link. However errors do accumulate, and in Fig. 9 the solid curve fits to the BER data at lower power levels show the expected error accumulation assuming that bit errors evolve in the way described earlier for concatenated AM-PSK transmission links with full differential precoding at the source. Also shown (dashed curves) is the linear accumulation of bit errors that would be expected in the case that high-speed recoding were applied at each node in the sequence of links. From these results we observe that the additional degradation in the overall BER due to full differential precoding at the source performs as expected and is bounded. In comparison with the approach of recoding at every node, the additional BER degradation is less than 1 order of magnitude after 10 links. This is in contrast to the other alternative approach of omitting all precoding and using receiver-based decoding, where error propagation is, in principle, unbounded.

As noted earlier, the programmed guard band between packets was 150 bit periods in duration, despite which no significant patterning was observed, as can be seen in Fig. 8. The maximum guard band that could be used is restricted by the low-frequency cut-off of components such as the DDAs. In addition, Fig. 8 shows the effects of launching data packets with different preprogrammed preamble patterns, resulting in different data patterns in the guard band when the DWSLs are switched. In the case of a preamble resulting in all zeros [as in Figs. 8(a) and 8(b)] or all ones, after one stage of conversion to the duobinary format the preamble is converted to all zeros (which appears in the traces as all ones because of the noninverting receiver used). This zero level is then maintained as the packet passes through each subsequent stage of format conversion. Figure 8(a) illustrates the pulse overshoot due to the switch-over of DWSL2, along with the trailing edge of the packet received from DSWL1, as described above. The trace recorded after eight circulations [Fig. 8(b)] shows that the trailing edge of the packet has been truncated by the recirculating loop and that the DWSL2 switch-over spike has been retained as a sequence of spurious

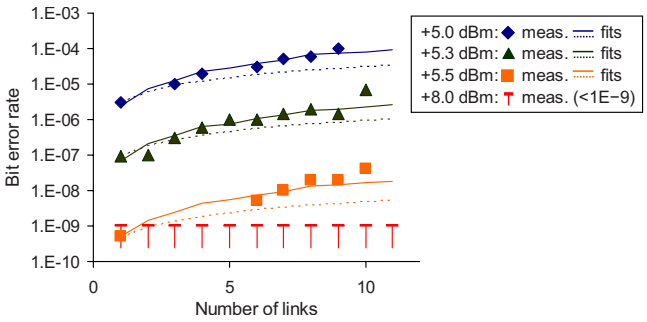

Fig. 9. Measured BERs of data packets versus number of concatenated links, for different powers incident on the receiver PD. Also shown are curve fits assuming full differential precoding at the packet source (solid curves) and precoding at every intermediate node (dashed curves).

bits. The width of this field of spurious bits increased by one on each circulation, as a consequence of the duobinary format conversion at each link. In effect, the switch-over spike introduced an isolated bit error, which then evolved and multiplied in the manner described earlier. In Figs. 8(c) and 8(d), the preamble is programmed so that the sequence 101010 is transmitted by MZM1 when DSWL1 is configured to switch a packet into the recirculating loop; in this case the preamble is converted during the first circulation of the loop to a sequence of all ones [equivalent to all zeros in the measured trace, Fig. 8(c)], and is converted after two or more circulations to a sequence of all zeros [all ones in measured traces, such as Fig. 8(d)]. We also note, in this case, that the spurious effects of the switch-over spike from DWSL2 are attenuated after one circulation and are entirely suppressed after two or more circulations, which could allow a reduction in the necessary guard band.

6. Discussion

It has been shown that the use of duobinary coding, in the form of AM-PSK optical modulation, allows narrower channel spacing in multiple concatenated links of a WDM-CR optical packet network. The narrower channel spacing allows a greater number of discrete channels to be accessed by the tunable laser transmitters in the network nodes, and this in turn leads to increased scalability of the network in terms of the total number of nodes. Appropriate differential duobinary precoding of data packets prior to insertion into the input buffers avoids the need to recode digitally at the full transmission line rate at each intermediate node, thus reducing the amount of high-speed electronics required. This approach of using full differential precoding at the source is suitable for optical packet networks based on WDM-CR because in such networks the whole transmission path for a packet (and hence the number of links passed) is exactly predetermined at the source.

A further advantage of full duobinary precoding at the source is that the packet header can be separately programmed so that a unique readily recognizable pattern appears in the header after a predetermined number of links have been passed, to indicate that the packet destination has been reached. This may simplify the process of recognizing packets that have reached their destination, and furthermore, the use of a common header pattern to denote that the destination has been reached would allow a generic design for the pattern recognition function at all nodes.

Finally, it may be possible to exploit the effect illustrated in Fig. 4, and observed experimentally as shown in Fig. 8(b), in which the width of the field of spurious bits in the preamble increased by one after each link, as a simple time-to-live mechanism. In this method, a marker bit would be deliberately introduced into the preamble of launched packets, and similar to the error propagation effects discussed earlier, this single marker bit would give rise to a field of marker bits after propagation through several links. Thus a packet that traverses an excessive number of links could be readily identified by the large width of this marker field and then purged from the network.

7. Conclusions

In this paper we have demonstrated that a WDM clockwork-routed network may be constructed using 42.6 Gbit/s AM-PSK formatted data packets, dual-wavelength switchable lasers, and WDM 2×2 interleavers with 50 GHz channel spacing. We have demonstrated propagation over at least 15 consecutive links. This link count enables transmission on an 8×8 toroidal network of 64 supernodes, while the potential for 50 GHz channel spacing enables scaling to over 32 wavelength pairs, giving a total of 2048 optical network nodes. We have proposed and demonstrated a simplification of the high-speed electronics required for multilink AM-PSK packet transmission, by fully precoding data packets at the source according to the predetermined number of hops to the destination. We described added beneficial features for optical packet routing, including simplified destination recognition and time-to-live mechanisms. Finally, we have investigated the error multiplication effects associated with full differential precoding at the source in concatenated AM-PSK transmission and verified its negligible impact on the overall system performance.

Acknowledgments

This work was supported by Science Foundation Ireland under grants 03/IN3/I560 and 06/IN/I969.

References

1. Q. Yang, K. Bergman, G. D. Hughes, and F. G. Johnson, "WDM packet routing for high-capacity data networks," J. Lightwave Technol. **19**, 1420–1426 (2001).
2. A. Shacham and K. Bergman, "A fully implemented 12×12 data vortex optical packet switching interconnection network," J. Lightwave Technol. **23**, 3066–3075 (2005).
3. A. Shacham and K. Bergman, "Optimising the performance of a data vortex interconnection network," J. Opt. Netw. **6**, 369–374 (2007).
4. R. Hemenway, R. Grzybowski, C. Minkenberg, and R. Luijten, "Optical packet-switched interconnect for supercomputer applications," J. Opt. Netw. **3**, 900–913 (2004).
5. C. Minkenberg, F. Abel, P. Muller, R. Krishnamurthy, M. Gusat, and B. R. Hemenway, "Control path implementation for a low-latency optical HPC switch," in *Proceedings of 13th Symposium on High Performance Interconnects* (IEEE, 2005), pp. 29–35.
6. E. Bravi and D. Cotter, "Low-latency optical network based on wavelength division multiplexed clockwork routing," Photonic Network Commun. **14**, 83–88 (2007).
7. E. Bravi and D. Cotter, "Optical packet-switched interconnect based on wavelength-division-multiplexed clockwork routing," J. Opt. Netw. **6**, 840–853 (2007).
8. E. Bravi and D. Cotter, "Traffic analysis of optical networks based on wavelength division multiplexed clockwork routing," in *Proceedings of IEEE International Conference on Communications* (IEEE, 2007), paper ONS8.2.
9. F. Chevalier, D. Cotter, and D. Harle, "A new packet routing strategy for ultra-fast photonic networks," in *Proceedings of IEEE Globecom* (IEEE, 1998), pp. 2321–2326.
10. A. Linder, "The duobinary technique for high speed data transmission," IEEE Trans. Commun. Electron. **82**, 214–218 (1963).
11. X. Gu and L. C. Blank, "10 Gb/s unrepeatered three-level optical transmission over 100 km of standard fiber," Electron. Lett. **29**, 2209–2211 (1993).
12. G. May, A. Solheim, and J. Conradi, "Extended 10 Gb/s fiber transmission distance at 1538 nm using a duobinary receiver," IEEE Photon. Technol. Lett. **6**, 648–650 (1994).
13. K. Yonenaga and S. Kuwano, "Dispersion-tolerant optical transmission system using duobinary transmitter and binary receiver," J. Lightwave Technol. **15**, 1530–1537 (1997).
14. T. Ono, Y. Yano, K. Fukuchi, T. Ito, H. Yamazaki, M. Yamaguchi, and K. Emura, "Characteristics of optical duobinary signals in terabit/s capacity, high-spectral efficiency WDM systems," J. Lightwave Technol. **16**, 788–797 (1998).
15. M. Duelk, J. Gripp, J. Simsarian, A. Bhardwaj, P. Bernasconi, M. Zirngbl, and O. Laznicka, "Fast packet routing in a 2.5 Tb/s optical switch fabric with 40 Gb/s duobinary signals at 0.8 b/s/Hz spectral efficiency," in *Optical Fiber Communications Conference, Technical Digest* (Optical Society of America, 2003), Vol. 3, paper PD08.
16. F. Smyth, E. Connolly, B. Roycroft, B. Corbett, P. Lambkin, and L. P. Barry, "Fast wavelength switching lasers using two-section slotted Fabry-Perot structures," IEEE Photon. Technol. Lett. **18**, 2105–2107 (2006).