# Multi-Wavelength Coding for Packet-Switched Optical Interconnects

Andrew West, Simon Moore
Computer Laboratory, University of Cambridge

Michael Dales, Madeleine Glick
Intel Research Cambridge

**Abstract:**

High-capacity optical networks using Wavelength Division Multiplexing (WDM) are well established in the long-haul network arena and with low-power, highly-integrated optical components continuing to reduce in cost, fibre's vast bandwidth advantages are ready to be deployed in demanding short-range networks. We employ data striping across the multiple wavelengths simultaneously, a method of bandwidth aggregation scalable to 100s of Gigabits/s per fibre. This differs from usual WDM where each wavelength represents a logically separate flow.

In this paper we exploit asynchronous coding to significantly increase throughput and reduce latency of multi-wavelength packet-switched interconnects. In addition, we outline an approach for two-dimensional constrained coding with a motivational example of controlling gain modulation of a Semiconductor Optical Amplifier (SOA), with overhead in the 10-20% range. Finally we briefly discuss the detection of line errors and their implications for these codes.

## 1. Introduction

Fibre optic communication is common for long distances, but is becoming attractive over shorter distances. Fibre has some significant advantages over copper, most importantly in terms of bandwidth availability, in system noise reduction, and power saving. In deployed wide area networks, WDM has been seen to achieve data rates of 1.6 Terabits [1].

Despite the success of optical interconnects on the large scale networks, optical networking has had less impact on the short-range interconnects, with copper still being the primary transmission medium. The continued reduction in cost of low-power, highly integrated optical components is starting to make them competitive with copper interconnects over shorter distances. Whilst an obvious route to optics is simply to replace copper with fibre, there is also the opportunity to take advantage of the advances made in optical network design in the long-haul. Many advances, such as WDM and optical switching, can be adapted for short-range interconnects. Other technologies are less relevant, for example chromatic dispersion compensation is unnecessary on such short distance links.

In long-haul WDM, each wavelength usually carries a logically separate flow of information and switching is performed with the use of per-wavelength add/drop multiplexers. We consider an packet-switched optical network in which data is transmitted with packet granularity, striping data over wavelengths in parallel to create a 'light-bus' [2]. In *Section 2*, we present a scheme using asynchronous coding which significantly reduces clock recovery synchronization cost when optically switching on a per-packet basis, allowing higher overall system throughput.

In *Section 3*, we look at the use of SOAs, devices which can act as an broad-wavelength-range amplifier and electrically-controlled optical switch. They offer nanosecond time-scale switching but their optical gain is subject to data-dependent modulation. To reduce the optical penalty and receive eye degradation, this section outlines an electrical means of encoding the data prior to transmission to avoid optical penalty, relaxing the optical power budget and potentially reducing cost, unlike work which seeks address the issue with additional optical equipment [3].

## 2. Reducing Synchronization Overhead

In a packet-switched optical network, electronic circuitry performs clock recovery on the output of every optical receiver to be able to sample incoming data on that wavelength at the correct time during each bit period. For a clock recovery circuit to determine the sampling position with a high degree of confidence, it is commonplace to allow 1000 bit periods or more for this lock acquisition. During this time, a high transition density must be maintained for the training of the decision circuit. The conventional approach is to precede every data frame with a preamble of 'idle' characters for long enough for the clock to be recovered. This method is clearly wasteful when optical switching may occur very frequently.

We propose the use of asynchronous coding to avoid the synchronization overhead of a preamble sequence by using channel coding that does not require clock recovery. This is achieved with an *m-of-n* data coding (Table 1) where valid codewords have exactly *m* bits set and *m* bits cleared on *n* wavelengths. For instance, with a *3-of-6* code we have 20 valid codewords and we assign 16 of them for data allowing us to convey 4 bits of information per codeword. The other 4 symbols are used for meta-data, similar to the provision of K-symbols in the 8B10B block code [4].

| *Number of wavelengths* | *Code* | *Typical payload data transmitted per bit period (bits)* | *Typical overhead (%)* |
|---|---|---|---|
| 4 | *2-of-4* | 2.5 | 60 |
| 6 | *3-of-6* | 4 | 50 |
| 8 | *4-of-8* | 6 | 34 |
| 10 | *5-of-10* | 7.5 | 34 |
| 16 | *8-of-16* | 13 | 24 |

**Table 1.** Properties of *m*-of-*n* coding schemes as employed in multi-wavelength coding.

These *m-of-n* codes are amongst those used in Delay Insensitive (DI) asynchronous circuit implementations [5], [6]. With the use of per-wavelength deskew circuitry, and a training pattern as outlined earlier, we can transfer a different codeword in each bit period. To ensure transition density, we never transfer the same codeword in two successive cycles but reserve at least one codeword meaning “treat as identical to previous word” and alternate/cycle as required.

Given the overhead for theses asynchronous codes are relatively high, we continue to recover the clock as usual and then migrate to a more efficient conventional block-based code once clock synchronization is complete. The advantage of our asynchronous coding is in making use of the preamble time to transfer useful information.

When we have a multi-wavelength interconnect, only a single clock recovery unit is required at the receiver as it is known the data bits on each of the wavelengths were referenced against the same transmit clock. Use of a single clock recovery circuit instead of one per wavelength may in itself provide a notable silicon area and power saving. At the electrical outputs from the optical receivers, there will be observed skew between the wavelengths which can be handled in an integrated circuit by using programmable delay lines. Sending a known 'training pattern' in the preamble allows correlation between the bitstreams received from the wavelengths to determine the relative skew. The circuitry can thereby set a per-wavelength delay that re-aligns the bitstreams received from each optical receiver.

The length of the preamble usually required for clock recovery, frame size and comparative overhead between asynchronous and synchronous codes will dictate the achievable increase in throughput (Figures 1 and 2). To take a specific example, with an 8-wavelength system with 25% overhead synchronous code (8B10B) using a clock recovery circuit taking 1000 bit periods to lock, the time taken to transfer a 1520 byte frame is reduced from 2900 bit periods to 1990 bit periods with the addition of asynchronous coding. Intuitively, longer frames provide less dramatic gains as the preamble is a smaller proportion of the overall transfer time.

### 2.1 Selection of Block-Based Code

The selection of conventional block-based code is important as the amount of overhead will reduce system throughput. On a system using 8B10B line coding, 25% of the bits transmitted are consumed i) allowing control encodings, ii) ensuring DC balance and iii) limiting run-length to 5 successive identical bits. At current line rates, such high proportions of overhead are untenable and led to the development of alternatives such as 64B66B (as employed in 10 Gigabit Ethernet [7]). With 64B66B, XOR-based scrambling statistically guarantees bounded digital sum variation of the data stream. Rather than the unified control/data character set of 8B10B, 64B66B employs separate frame formats for data, control and mixture.

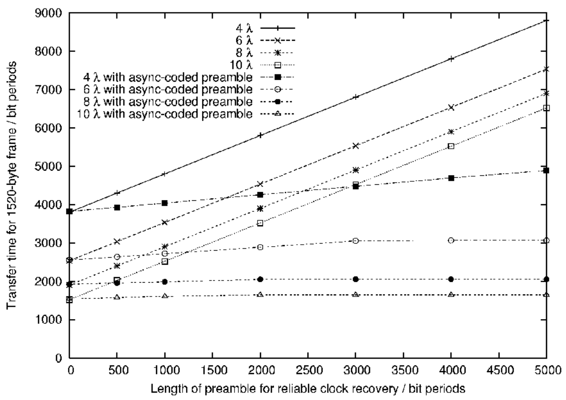


**Figure 1.** Asynchronous coding during preamble significantly reduces transfer time when clock recovery is relatively slow in a packet-switched optical link.

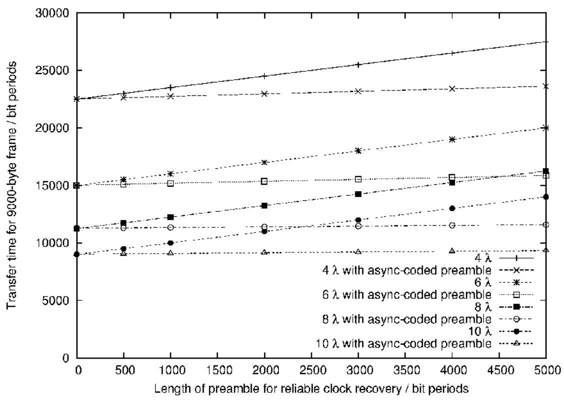


**Figure 2.** Increasing the payload size to a 'Jumbo' Ethernet frame shows a modest throughput increase with asynchronously-coded preambles.

We can devise good codes for multi-wavelength use by integrating the most attractive features consistent with the short-range packet-switched optics environment. Ideal features include a low overhead, a structure with a short internal block length to allow rapid alignment after reset (like 8B10B), separate control/data frame types, relatively short run-length. Specifying a short run-length, such as 4 to 6 identical bits, is useful as it limits the time a clock recovery circuit must remain internally locked.

## 3. Coding to Average Optical Power in SOAs

As an example of developing more constrained multi-wavelength coding, we will continue to examine how to reduce the pattern-dependent gain modulation of an SOA. We *temporally* apply run-length constraints on a per-wavelength basis and limit the number of simultaneous '1' bits *across* the wavelengths in any given bit period. There are other situations where similar types of two-dimensional constrained codes are worthy of consideration. An example in the optical domain would be where crosstalk between multiple adjacent wavelengths could be compensated for by adding additional constraints to avoid certain codewords or groupings that would otherwise exacerbate the situation and increase the likelihood of bit errors.

We consider a *block* as being a fixed number of bit periods across $n$ wavelengths. To meet our run-length constraint of $r$ bits, we set the block length to be $r$ bits long. We define a minimum number of simultaneously-on wavelengths as $j$ and a maximum as $k$. When $j = k$ an SOA is operating in an optimal manner as this will maintain a relatively consist number of charge carriers and will suppress the cause of gain modulation, this also gives us coding that is very similar to the asynchronous style presented earlier, however, with a guaranteed maximum run-length. By relaxing the bounds $j$ and $k$ we can control the variance on the number of wavelengths simultaneously '1', less demanding bounds will allow lower overhead code implementations. On each wavelength we restrict the codes to to those that do not violate the run-length constraint: if the previous bit in the last block on that wavelength was '0' we can use any value but $r$ '0's, if the previous bit on that wavelength was '1' we encode with a value but $r$ '1's. By enumeration we can gain an idea of how costly the constraints are to meet. As an example, given ($n$, $j$, $k$, $r$) = (8, 3, 5, 5) we have a possible 2,138,647 valid blocks ($>2^{21}$).

We partition our input data into block-sized vectors in an buffer that distributes a number of input bytes over a number of blocks. In the case of control characters (K-symbols), such a character is expanded to consume a whole block. Doing so allows redundancy such that even a few bits in error will not translate a valid control character into a different control character or a data block. Considering the symmetry and permutation between the valid blocks, what may appear to be a huge state space can be reduced to a number of look-up tables and permutation logic in the encoder and decoder. The size and overall complexity depends greatly on the constraints. A design that uses SRAM to store tables may be reconfigured on the fly for various constraints, to achieve the desired optical link budget against bandwidth.

To summarize, coding to satisfy multiple constraints is non-trivial and depending on the application-specific definition of acceptable complexity, low overhead is possible. Coding with a block size $r = 3$ and $r = 5$ is shown in Figures 3 and 4 respectively. The lowest overhead with the constraints $(n, j, k, r) = (8, 3, 5, 5)$ is ~11%, with typical overheads ranging from 10-20%. Current on-going work is the hardware implementation of a dynamically configurable encoder and decoder.

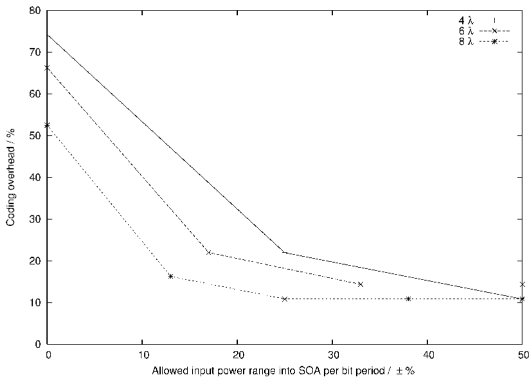


**Figure 3.** Multi-wavelength coding with a block size of 3 bounds SOA power with a moderate cost in overhead.

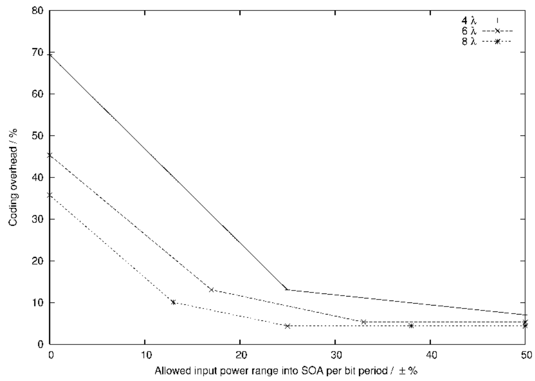


**Figure 4.** Multi-wavelength coding with a block size of 5 allows tight bounds on the average SOA input power with low overhead but far greater complexity of implementation.

## 4. Data Integrity

One of the potential pitfalls is that errors should be flagged reliably and the mean time to false packet acceptance should be of a satisfactory order of magnitude. A single bit error will likely cause an avalanche of subsequent bit modifications in the resulting data payload. This contrasts poorly with simpler single-channel code schemes such as 8B10B where errors propagate a limited distance. The CRC protecting the payload may allow some burst errors to go undetected. It is our suggestion that a compromise between complexity and reliability is to include circuitry in the encoder and decoder that performs a cryptographic hash generation and verification. This would offer increased confidence that corrupted data would not be accepted by the host (depending on the hash length).

## 5. Conclusions

This paper has introduced a novel way for optical packet-switched multi-wavelength interconnects to reduce the clock synchronization cost by transmitting useful payload information in an asynchronous preamble. This is effective at significantly decreasing frame transfer time, improving overall throughput.

An example of multi-wavelength coding to mitigate SOA gain modulation has been presented, illustrating coding with two-dimensional constraints. Development of these techniques should lead to parametric, efficient codes to flexibly optimize these and other optical interconnect systems.

## References


[1] A. R. Pratt, P. Harper, S. B. Alleston, P. Bontemps, B. Charbonnier, W. Forysiak, L. Gleeson, D. S. Govan, G. L. Jones, D. Nesset, J. H. B. Nijhof, I. D. Phillips, M. F. C. Stephens, A. P. Walsh, T. Widdowson, and N. J. Doran. 5,745 km DWDM Transcontinental Field Trial Using 10 Gbit/s Dispersion Managed Solitons and Dynamic Gain Equalization. In *Optical Fiber Communications Conference*, pages 1–3, March 2003.

[2] M. R. Loeb and G. R. Stilwell. High Speed Data Tansmission on an Optical Fibre Using a Byte Wide WDM System. *Journal of Lightwave Technology*, 6:1306, 1988.

[3] Hyang K. Kim and S. Chandrasekhar. Reduction of Cross-Gain Modulation in the Semiconductor Optical Amplifier by Using Wavelength Modulated Signal. *IEEE Photonics Technology Letters*, Vol. 12, No. 10, October 2000.

[4] A. X. Widmer and P. A. Franaszek, “A DC-Balanced, Partitioned-Block, 8B/10B Transmission Code,” *IBM Journal of Research and Development*, vol. 27, no. 5, pp. 440–451, Sept. 1983.

[5] John Bainbridge and Steve Furber. CHAIN: A delay-insensitive chip area interconnect. *IEEE Micro*, 22:16–23, 2002.

[6] Jens Sparso and Steve Furber. *Principles of asynchronous circuit design: a systems perspective*. Kluwer Academic Publishers, ISBN 0792376137, 2001.

[7] IEEE, “IEEE 802.3ae — 10 Gb/s Ethernet,” 2002, standard.