Combining Forward Error Correction and Network Coding in Bufferless Networks: a Case Study for Optical Packet Switching

Gergely Biczók*‡, Yanling Chen†‡, Katina Kralevska‡ and Harald Øverby‡
Email: biczok@tmit.bme.hu, yanling.chen@uni-due.de, katinak@item.ntnu.no, haraldov@item.ntnu.no
*MTA-BME Future Internet RG, Budapest University of Technology and Economics
†Institute of Digital Signal Processing, University of Duisburg-Essen
‡Dept. of Telematics, Norwegian University of Science and Technology

***Abstract*—Bufferless network operation is favorable in many application domains such as industrial networks, on-chip networks and optical packet switching (OPS). The main challenge with zero buffers is the avoidance or handling of contention; indeed, many domain-specific contention resolution techniques have been proposed in the literature. In this paper, we propose a generic combined forward error correction (FEC) and network coding (NC) scheme, which mitigates the negative impact of contentions at the network layer. Specifically, we present a case study for OPS utilizing FEC at the ingress node and NC at an intermediary optical packet switch to reduce packet loss due to contention. Our analysis shows that if used in a smart way, our mechanism can reduce decoding error and packet loss with multiple orders of magnitude while adhering to buffering limitations and meeting delay requirements. We believe that such a combined coding scheme has the potential to be utilized both in OPS (data center and core networks) and other networks where (near-)zero buffers are required.**

I. Introduction

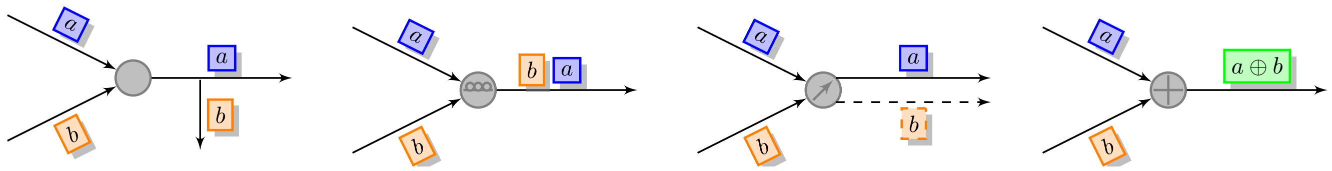
Recently, networking technology has become a key ingredient in many application domains. Each application domain comes with its own set of functional and non-functional requirements, such as ultra-low delay for real-time systems or energy efficient operation for embedded systems. Interestingly, the performance requirements of several application domains can be satisfied by bufferless operation, i.e., by using (near)-zero buffers in the forwarding elements of the corresponding networks. In turn, bufferless operation requires either avoiding or handling contention events, when multiple packets are aligned to be forwarded by the same element.

For example, wireless industrial automation networks operate in synchronous mode (for deterministic data delivery) and in a multi-hop fashion (as these networks are large) with strict delay requirements between sensor and actuator. A bufferless design can satisfy these requirements, if appropriate system-wide transmission scheduling is applied to avoid contention events at in-network switches [1]. Furthermore, on-chip networks, connecting multiple cores in an integrated circuit, have energy consumption, spatial density and complexity constraints in addition to requiring ultra-low communication delay [2]. Techniques such as contention-aware scheduling [3] and specialized flow control [4] have been proposed and studied to enable efficient bufferless operation. Finally, Optical Packet Switching (OPS), a potential all-optical networking architecture enabling statistical multiplexing, requires a very low packet loss rate while avoiding the use of expensive optical buffering [5]. Therefore, several complicated contention avoidance and resolution techniques have been proposed [6]. We believe that all of these application domains can benefit from a unified approach to contention handling by means of a smart coding scheme. Throughout this paper, we focus on the OPS domain; however, our overall design and high-level findings could potentially be transferred to some or all of the above-mentioned application domains.

OPS has become a resurgent topic in the research community. One reason is the rise of interest in energy efficient communications [7]. If implemented with all-optical contention resolution, completely eliminating store-and-forward operation, OPS has the potential to reduce overall network energy consumption by an order of magnitude. Another strong driver behind OPS is the emergence of cloud computing and the underlying virtual and physical data center infrastructure. OPS is able to enhance the performance of data center networks by providing large-capacity and fast switching capability [8]. When combined with the Software-Defined Networking (SDN) paradigm (where a slower control plane is decoupled from a faster data plane), OPS [9] and hybrid optical packet/circuit switching [10] are both shown to deliver high performance. Still, optical contention resolution is quite challenging to implement in a cost efficient manner, while keeping packet loss rates low. Several techniques have been proposed in order to deal with such packet loss, e.g., the use of wavelength conversion, fiber-delay line buffering and deflection routing [6]. Wavelength conversion is efficient but potentially expensive, fiber-delay lines try to mimic electronic buffering, and are therefore costly and of limited use, while the performance of deflection routing largely depends on the network topology.

Putting these specific contention resolution mechanisms aside, packet loss in a communication network is very often alleviated by smart coding techniques. For example, in the context of OPS, the Network Layer Packet Redundancy Scheme

978-1-4799-8950-8/16/$31.00 ©2016 IEEE



(a) Contention w/ random dropping (b) Contention resolution w/ FDL (c) Contention resolution w/ wavelength conversion (d) Contention resolution w/ NC

Fig. 1. Contention resolution in OPS

(NLPRS) utilizes Forward Error Correction (FEC) at an OPS ingress router [11] to minimize packet loss. This mechanism largely increases the probability of successful packet reception at the respective egress router at the cost of some extra traffic inside the optical network. Another coding technique for mitigating packet loss proposed by many researchers for wireless networks is network coding (NC), e.g. [12]. By combining packets using NC (through an NC shim between the MAC and IP layers), packet loss (induced by the wireless medium) can be greatly reduced. A straightforward application of the above network coding technique is not feasible in OPS due to the lack of store-and-forward capabilities; however, with OPS being a loss system at the network level, the idea of combining packets which would otherwise be dropped is intriguing.

In this paper, we combine forward error correction and network coding in order to achieve very low packet loss rates in all-optical packet-switched networks. Rather than coupling FEC together with a randomly chosen contention resolution technique, we actually build on the redundancy introduced by FEC when applying network coding at the optical packet switch. By carefully selecting contention events at which we code the colliding packets together instead of dropping one of them, NC can transfer this redundancy across different packet flows. This both increases the probability of successful decoding and achieves a lower data packet loss rate. Our main contribution is twofold: (i) we introduce the design of the combined coding scheme and (ii) we model the case of a single optical packet switch as a bufferless network coding system, and provide an analytical characterization of the decoding probability and the data packet loss rate. Note that implementation details are out of scope for this paper; nevertheless, we do provide a discussion on the limitations of our work.

The rest of the paper is organized as follows. Section II describes the proposed combined forward error correction and network coding framework. Section III presents the analytical model and the resulting decoding probabilities and packet loss rates. We provide a discussion on limitations and future work in Section IV. Finally, Section V concludes the paper.

II. Combining FEC and Network Coding

Combining channel coding and physical network coding has been proposed for wireless networks [13], where the broadcast nature of the wireless medium can be exploited to boost throughput. In case of OPS networks on the other hand, the combination of the two coding paradigms can help reduce packet loss. We combine forward error correction at the edge of the optical network with network coding in the core optical packet switches. Note that another type of FEC is used in optical transmission at very high speeds in connection with advanced modulation schemes [14]. Moreover, the use of NC has been proposed at the optical layer for dedicated protection [15]. In this paper, we utilize FEC at the network layer operating on data packets, and NC inside optical packet switches.

A. *Forward error correction at the edge*

FEC is widely used in communication systems to ensure successful transmission over unreliable channels. We consider utilizing a Maximum Distance Separable (MDS) (n, k) linear block code, where every k units of information are encoded into n units of information, which are linear combinations of the k original units. The redundant $r = n - k$ units are responsible for error control; in our case this means alleviating packet loss due to contentions in OPS. Intuitively, to be able to recover the k original information units, any k out of n coded units are needed for decoding at the destination. MDS codes reach the Singleton-bound [16]. The popular Reed-Solomon (RS) codes belong to the MDS class; throughout this paper, we assume systematic RS codes in line with [17] when mentioning FEC.

FEC encoding and decoding take place at the electronic ingress and egress router at the edge of the OPS network, similarly to NLPRS [11]. Data packets arrive from access networks to an electronic ingress router of the OPS network. Upon arrival, packets are grouped according to their access network destination. Packets with the same destination access network (i.e., egress router) form a packet set of size k (k being a small integer). The k packets in a set are encoded into n packets by adding r redundancy packets via FEC. All packets should have the same size that is at least the size of the largest data packet in the packet set.

FEC is shown to reduce packet loss rates by orders of magnitude in certain scenarios [11]. However, a low code rate (k/n) implying high overhead (r/k) introduces unfavorable burstiness in addition to decreased resource utilization, calling for smarter techniques. On the one hand, electronic FEC can be combined with any contention resolution scheme operating in the OPS core. On the other hand, choosing a contention resolution scheme which actually builds on the strengths of FEC is highly desirable.

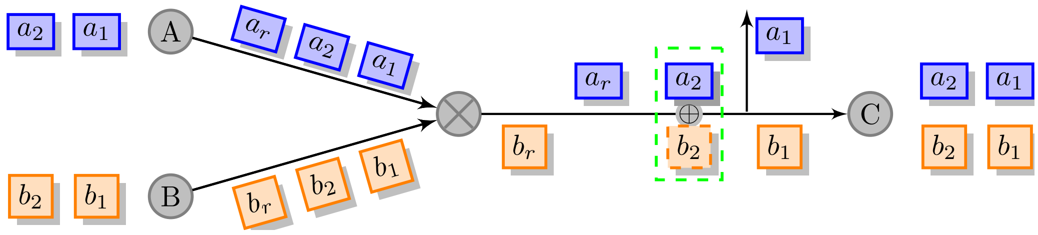


Fig. 2. FEC combined with NC. A, B and C are ingress/egress routers. Two packet sets from different ingress nodes contend at the optical switch.

B. Adding contention resolution: network coding

Contention resolution is a central topic in optical packet switching: contention can occur when multiple packets are assigned to the same output wavelength. Since it is very unlikely in a properly designed optical packet switch that more than 2 packets contend for the same wavelength at the same time, we focus on the 2-packet case. The default behavior in this case is to drop one of the packets randomly (see Fig. 1(a)). If contention is not handled properly, packet loss rate inside the optical network can be prohibitively high. Therefore, several techniques have been proposed in the literature. Established methods for handling contentions include Fiber Delay Line (FDL) buffering (Fig. 1(b)) and wavelength conversion (Fig. 1(c)). Both of these techniques work well, but (i) they are quite costly in terms of extra hardware, and (ii) they operate independently from the FEC mechanism.

We propose using simple XOR network coding for contention resolution in the optical packet switch (see Fig. 1(d)). Note that optical XOR logic exists, enabling packets to stay in the optical domain [18]. Although not providing contention resolution in itself, NC can build on the redundancy in packet sets introduced by the FEC. Specifically, NC can transfer the redundancy inherent in an encoded packet set to the contending packet set by binding the packet sets together. This way, the original information in both packets sets can be recovered with higher probability at the corresponding egress router. There are two main requirements in terms of successful decoding at the egress. First, we only XOR packets if their corresponding packet sets are headed towards the same egress router. Note that this is not guaranteed for every contending packet set pair: being switched to the same wavelength generally means only a common next hop. Second, at least k packets from one of the packet sets have to be transmitted successfully. The resulting joint error-correcting/network coding scheme can be seen in Fig. 2.

An example with multiple contentions and a $(3, 2)$ MDS code is provided in Fig. 2. Assume that packets from two sources A and B, (two different ingress nodes) but with the same destination egress, contend at an OPS core node. Packet a_1 is discarded, however, packet $a_2 \oplus b_2$ is transmitted instead of discarding either a_2 or b_2. Thus, at the OPS egress node, packet b_2 is reconstructed using the remaining packets from sender B; a_2 is reconstructed from b_2 and $a_2 \oplus b_2$; and finally a_1 is reconstructed from the packet a_r and the newly reconstructed a_2. Reconstruction would not be possible if a_2 was dropped instead of being intercoded with b_2.

Since the utilized FEC is indeed a systematic RS code, the actual decoding process depends on which type of packets have been successfully received and XOR-ed, respectively. Using the example in Fig. 2, there are two different cases:

- We have a_r, b_1, b_r and $a_2 \oplus b_2$: we first RS-decode to get b_2, then XOR $a_2 \oplus b_2 \oplus b_2 = a_2$, then RS-decode to get a_1. In total, we have to perform two RS-decoding and one XOR operation to recover all original information.
- On the contrary, if we would have a_2, b_1, b_2 and $a_r \oplus b_r$: we first RS-encode to get b_r, then XOR $a_r \oplus b_r \oplus b_r = a_r$, then RS-decode to get a_1. In total, we have to perform one RS-encoding, one RS-decoding and one XOR operation.

Network coding enables redundancy to carry over from one packet set into the other by XORing them together. If there are multiple contentions inside a packet set, it is not trivial which combination of XORing and random dropping (which is the default operating mode of an optical packet switch) maximizes the available information for decoding at the egress. We expect that the best strategy largely depends on the value of contention probability and the FEC code parameters; we analyze this issue in Section III.

Furthermore, when considering packet sets traveling across multiple packet switches inside an OPS domain, it is not straightforward to ensure an optimal mixture of coding and dropping inside a packet set. Indeed, designing a coordination mechanism among packet switches constitutes important future work for us.

III. Analysis

Here we provide an analysis of our coding scheme as used in a simplified setting, where all packets pass through exactly one optical packet switch. We derive the condition under which our scheme provides a higher probability of successful decoding than bare FEC. Furthermore, we characterize the resulting data packet loss rates, which in turn determines the resource efficiency of the system.

A. Decoding probability

Assume that packets from two different ingress nodes arrive at the same destination. We denote the two sources to be A and B. Packets may collide at an OPS core node. When contention occurs, for instance packet a_i and packet b_i are aligned to the same output wavelength, by random dropping either a_i or b_i will be randomly discarded; while by XOR, packet $a_i \oplus b_i$ is transmitted. Furthermore, suppose that

- the optical packet switch has 0 buffer;

- the transmission is slotted and synchronized;
- contention probability is $p < 1/2$ for each time slot;
- A and B use the same FEC;
- a header describing the information on the contentions is sent to the decoder by using a separate, faster and secure channel.

Let us consider the transmission of coded source A and source B, i.e., packet sets $a = a_1 \dots a_n$ and $b = b_1 \dots b_n$, through OPS. In order to evaluate the performance of the system by the probability of unsuccessful decoding, we consider the following two scenarios: 1) only random dropping; 2) first XOR then random dropping. Our objective is to find out the best strategy to use XOR and random dropping instead of random dropping only, so that the probability of unsuccessful decoding at the destination is minimized. Note that it is sufficient to analyze the probability of unsuccessful decoding of a only, since b has the same unsuccessful decoding probability due to the symmetric roles they play in the transmission.

Denote the number of the contentions as n_c; the number of XORs as n_x, the number of the randomly dropped packets from the coded packet sets a and b as n_c^a and n_c^b, respectively. Clearly we have $0 \leq n_c \leq n$ and $0 \leq n_c^a, n_c^b \leq n_c$. Recall that for each i ($1 \leq i \leq n$) contention occurs with a probability p, and a_i is dropped randomly with a probability $1/2$. In the case of $n_x = 0$, we have

$$\Pr(n_c = t) = \binom{n}{t} p^t (1-p)^{n-t};$$

$$\Pr(n_c^a = s | n_c = t) = \binom{t}{s} \left(\frac{1}{2}\right)^t.$$

1) FEC with random dropping: By the random-dropping strategy, when a contention occurs to packets a_i and b_i, either a_i or b_i is randomly discarded. In this case, the unsuccessful decoding for a happens as the following event occurs:

- $\mathcal{E}$: less than k packets from the coded source A are received at the destination, i.e., more than r from n packets are discarded during the transmission.

Event $\mathcal{E}$ takes place as $n_c^a \geq r + 1$. So we have the probability of unsuccessful decoding for a to be p_r and

$$p_r = \Pr(\mathcal{E}) = \sum_{t=r+1}^{n} \Pr(n_c = t) \sum_{s=r+1}^{t} \Pr(n_c^a = s | n_c = t)$$
$$= \sum_{t=r+1}^{n} \binom{n}{t} p^t (1-p)^{n-t} \sum_{s=r+1}^{t} \binom{t}{s} \left(\frac{1}{2}\right)^t.$$

2) FEC with first-XOR-then-random dropping: In the extended system, the first approach we take is to employ XOR at the first contention; and apply random dropping to the subsequent contentions. In this case, the unsuccessful decoding for a happens as the following events occur:

- $\mathcal{E}_1$: more than r from n packets are discarded during the transmission;
- $\mathcal{E}_2$: $k-1$ packets from a are received at the destination, whilst the knowledge of b could not help to recover a successfully.

Then the probability of unsuccessful decoding for a in this case is p_{x+r} by

$$p_{x+r} = \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2).$$

Events $\mathcal{E}_1$ and $\mathcal{E}_2$ take place as the XOR is employed at the first contention. So we have in this case $n_x = 1$ and

$$\Pr(n_x = 1, n_c^a = s | n_c = t) = \binom{t-1}{s} \left(\frac{1}{2}\right)^{t-1}.$$

In particular, event $\mathcal{E}_1$ occurs as $n_c^a \geq r + 1$. Therefore,

$$\Pr(\mathcal{E}_1)$$
$$= \sum_{t=r+2}^{n} \Pr(n_c = t) \sum_{s=r+1}^{t-1} \Pr(n_x = 1, n_c^a = s | n_c = t)$$
$$= \sum_{t=r+2}^{n} \binom{n}{t} p^t (1-p)^{n-t} \sum_{s=r+1}^{t-1} \binom{t-1}{s} \left(\frac{1}{2}\right)^{t-1}.$$

Event $\mathcal{E}_2$ occurs as $n_x = 1, n_c^a = r$, and knowledge of b at the destination could not help to recover a. We note that if $n_c^b \leq r - 1$, then at least k of n packets from source B will be received at the destination, and thus, b will be decoded successfully. Therefore, knowledge of b will clear the confusion of the XOR and further help a to obtain one extra packet, which in addition to another $k-1$ packets from a, is sufficient to have a recovered successfully. So event $\mathcal{E}_2$ takes place only if $n_c^b \geq r$. Recall that $n_x = 1, n_c^a = r$. Thus we have $n_c = n_x + n_c^a + n_c^b \geq 2r + 1$. So

$$\Pr(\mathcal{E}_2) = \sum_{t=2r+1}^{n} \Pr(n_c = t) \Pr(n_x = 1, n_c^a = r | n_c = t)$$
$$= \sum_{t=2r+1}^{n} \binom{n}{t} p^t (1-p)^{n-t} \binom{t-1}{r} \left(\frac{1}{2}\right)^{t-1}.$$

Note that in case of the sufficient redundancy such that $r \geq n/2$, we have $\Pr(\mathcal{E}_2) = 0$. Easily we have:

Theorem 1: First-XOR-then-random dropping is better than random dropping in the terms of reducing the unsuccessful decoding probability if $r \geq \lceil (n-1)/2 \rceil$.

Proof: We can divide p_r into two parts, i.e., $p_r = p_r^1 + p_r^2$,

$$p_r^1 = \sum_{t=r+1}^{n-1} \binom{n}{t} p^t (1-p)^{n-t} \sum_{s=r+1}^{t} \binom{t}{s} \left(\frac{1}{2}\right)^t;$$

$$p_r^2 = \left(\frac{p}{2}\right)^n \sum_{s=r+1}^{n} \binom{n}{s}.$$

In case of $r \geq \lceil n/2 \rceil$, we have $r \geq n/2$ and thus $\Pr(\mathcal{E}_2) = 0$. We show that $p_{x+r} < p_r$ as follows:

$$\begin{aligned} p_{x+r} &= \Pr(\mathcal{E}_1) \\ &= \sum_{t=r+2}^{n} \binom{n}{t} p^t (1-p)^{n-t} \sum_{s=r+1}^{t-1} \binom{t-1}{s} \left(\frac{1}{2}\right)^{t-1} \\ &= \sum_{t=r+1}^{n-1} \binom{n}{t+1} p^{t+1} (1-p)^{n-t-1} \sum_{s=r+1}^{t} \binom{t}{s} \left(\frac{1}{2}\right)^{t} \\ &\overset{(*)}{<} \sum_{t=r+1}^{n-1} \binom{n}{t} p^t (1-p)^{n-t} \sum_{s=r+1}^{t} \binom{t}{s} \left(\frac{1}{2}\right)^{t} \\ &= p_r^1 < p_r, \end{aligned}$$

where (*) holds due to the fact that $p \leq 1/2$; and $\binom{n}{t+1} < \binom{n}{t}$ as $t \geq r+1 \geq (n+1)/2$.

In case of $r = (n-1)/2$ for an odd n, $\Pr(\mathcal{E}_2) > 0$. Firstly, from above proof, we have $\Pr(\mathcal{E}_1) < p_r^1$. To prove $p_{x+r} < p_r$, we just need to show that $\Pr(\mathcal{E}_2) \leq p_r^2$. This is given by

$$\begin{aligned} \Pr(\mathcal{E}_2) &= 2\binom{n-1}{(n-1)/2} \left(\frac{p}{2}\right)^n \overset{(**)}{\leq} \sum_{s=(n+1)/2}^{n} \binom{n}{s} \left(\frac{p}{2}\right)^n \\ &= p_r^2, \end{aligned}$$

where (**) holds due to the fact that $2\binom{n-1}{(n-1)/2} \leq 2^{n-1} = \sum_{s=(n+1)/2}^{n} \binom{n}{s}$ for an odd $n \geq 3$. ■

3) FEC with n_0-XOR-then-random dropping: A more general approach which can be taken is to execute XOR at the first $n_0 > 1$ contentions and then random-dropping at the subsequent ones. In this case, the unsuccessful decoding for a happens only as the following event occurs:

- $\mathcal{E}_1^{n_0}$: more than r from n packets are discarded during the transmission;
- $\mathcal{E}_2^{n_0}$: $\geq k - n_0$ but $\leq k-1$ packets from a are received at the destination, whilst the knowledge of b could not help to recover a successfully.

It is easy to see that event $\mathcal{E}_1^{n_0}$ occurs as $n_c^a \geq r+1$ and $n_c \geq r + n_0 + 1$. Thus we have:

$$\begin{aligned} &\Pr(\mathcal{E}_1^{n_0}) \\ &= \sum_{t=r+n_0+1}^{n} \binom{n}{t} p^t (1-p)^{n-t} \cdot \sum_{s=r+1}^{t-n_0} \binom{t-n_0}{s} \left(\frac{1}{2}\right)^{t-n_0} \end{aligned}$$

and

$$\begin{aligned} &\Pr(\mathcal{E}_2^{n_0}) \\ &= \sum_{t=2r-n_0+2}^{n} \binom{n}{t} p^t (1-p)^{n-t} \cdot \sum_{s=r-n_0+1}^{r} \binom{t-n_0}{s} \left(\frac{1}{2}\right)^{t-n_0}. \end{aligned}$$

Note that in case of $2r - n_0 + 2 > n$, we have $\Pr(\mathcal{E}_2^{n_0}) = 0$. In particular, we have:

Theorem 2: n_0-XOR-then-random dropping is better than (n_0-1)-XOR-then-random dropping as long as $n_0 < 2(r+1) - n$.

Proof: It is due to the facts that as $n_0 < 2(r+1) - n$, $\Pr(\mathcal{E}_2^{n_0}) = \Pr(\mathcal{E}_2^{n_0-1}) = 0$ and $\Pr(\mathcal{E}_1^{n_0}) < \Pr(\mathcal{E}_1^{n_0-1})$ by

$$\begin{aligned} &\Pr(\mathcal{E}_1^{n_0}) \\ &= \sum_{t=r+n_0+1}^{n} \binom{n}{t} p^t (1-p)^{n-t} \cdot \sum_{s=r+1}^{t-n_0} \binom{t-n_0}{s} \left(\frac{1}{2}\right)^{t-n_0} \\ &= \sum_{t=r+n_0}^{n-1} \binom{n}{t+1} p^{t+1} (1-p)^{n-t-1} \\ &\quad \cdot \sum_{s=r+1}^{t-n_0+1} \binom{t-n_0+1}{s} \left(\frac{1}{2}\right)^{t-n_0+1} \\ &< \sum_{t=r+n_0}^{n-1} \binom{n}{t} p^t (1-p)^{n-t} \cdot \sum_{s=r+1}^{t-n_0+1} \binom{t-n_0+1}{s} \left(\frac{1}{2}\right)^{t-n_0+1} \\ &< \sum_{t=r+n_0}^{n} \binom{n}{t} p^t (1-p)^{n-t} \cdot \sum_{s=r+1}^{t-n_0+1} \binom{t-n_0+1}{s} \left(\frac{1}{2}\right)^{t-n_0+1} \\ &= \Pr(\mathcal{E}_1^{n_0-1}). \end{aligned}$$

This completes our proof. ■

B. Packet loss rate

Now, we derive the data packet loss rates (PLR) for FEC with random dropping and our proposed combined FEC/NC scheme with 1 or n_0 XORs. Denote the total number of packets sent from a coded source as n where the number of data packets (DPs) is k and the number of redundancy packets (RPs) is r. We use Table I to examine the number of lost DPs when different schemes are used. Since lost DPs can be reconstructed from successful arrivals of DPs and RPs, the number of lost DPs may decrease after a possible reconstruction. We denote the lost DPs after reconstruction as DPAR. Note that lost DPAR $\leq$ lost DPs. If the total number of lost DPs (i) and lost RPs (j) in a packet set is greater than the total number of transmitted RPs (r), then reconstruction is not possible and the number of lost DPAR equals the number of lost DPs.

1) FEC with random dropping: When contention occurs to packet a_i and b_i, either a_i or b_i is randomly discarded. Thus, the maximum number of lost DPs in a single pair of packet sets is k, as at every contention event, exactly one of the colliding packets is dropped. A packet is lost when the number of contentions inside a packet set is $t > r$. The average number of lost DPs when t contentions happen inside a single packet set is

$$\begin{aligned} PLR &= \frac{1}{2k} \sum_{t=r+1}^{n} \left(\binom{n}{t} p^t (1-p)^{n-t} \frac{\sum_{i=t-r}^{\min(k,t)} i \binom{k}{i} \binom{r}{t-i}}{\binom{n}{t}} \right) \\ &= \frac{1}{2k} \sum_{t=r+1}^{n} \left(p^t (1-p)^{n-t} \sum_{i=t-r}^{\min(k,t)} i \binom{k}{i} \binom{r}{t-i} \right). \end{aligned}$$

2) FEC with first-XOR-then-random dropping: As it is shown in Table I, FEC with first-XOR-then-random dropping can alleviate one loss per packet set when the knowledge of the packets from the other packet set can help to recover

Lost DPs (i)	Lost RPs (j)	Lost DPAR with FEC	Lost DPAR with FEC+1 XOR	Lost DPAR with FEC+n_0 XORs
i=1	$0 \leq j \leq r-1$	0	0	0
	$j = r$	1	0 (*)	0 (*)
i=2	$0 \leq j \leq r-2$	0	0	0
	$j = r-1$	2	0 (*)	0 (*)
	$j = r$	2	1 (*)	0 (*)
$3 \leq i \leq k-1$	$0 \leq j \leq r-i$	0	0	0
	$j = r-i+1$	i	0 (*)	0 (*)
	$r-i+2 \leq j \leq r$	i	$i-1$ (*)	0 (*)
$i = k$	$0 \leq j \leq r-k$	0	0	0
	$j = r-k+1$	k	0 (*)	0 (*)
	$r-k+2 \leq j \leq r$	k	$k-1$ (*)	0 (*)

(*) packets from the other packet set can help recover the lost packets successfully

TABLE I
THE NUMBER OF LOST DPAR IN A PACKET SET AS A FUNCTION OF THE NUMBER OF LOST DPS AND RPS

the lost packets successfully. We always employ XOR at the first contention of DPs. A packet is lost when the number of contentions inside a packet set is $t > r+1$. In addition, we should consider that an extra packet is lost when the XORed packet cannot be decoded (two lost packets due to XORing at the first contention). Note, that this extra packet loss is far more unlikely to happen than adequate decoding, since $p^r >> p^{r+1}$ with a small p.

In particular, the average data packet loss rate is:

$$PLR = \frac{1}{2k}\sum_{t=r+2}^{n}\binom{n}{t}p^t(1-p)^{n-t} \cdot \frac{\sum_{i=t-r-1}^{\min(k-1,t-1)}(i+2)\binom{k-1}{i}\binom{r}{t-i-1}}{\binom{n-1}{t-1}}$$
$$= \frac{1}{2k}\sum_{t=r+2}^{n}\frac{n}{t}p^t(1-p)^{n-t} \cdot \sum_{i=t-r-1}^{\min(k-1,t-1)}(i+2)\binom{k-1}{i}\binom{r}{t-i-1}.$$

3) FEC with n_0-XOR-then-random dropping: A more general approach is taken when the first $n_0 > 1$ contentions of DPs are XORed, and then packets in subsequent contentions are randomly dropped. A packet is lost when the number of contentions inside a packet set is $t > r+n_0$. Also in this case, extra packets can be lost when the XORed packets cannot be decoded ($2n_0$ lost packets at n_0 XORed contentions).

The average data packet loss rate is:

$$PLR = \frac{1}{2k}\sum_{t=r+n_0+1}^{n}\binom{n}{t}p^t(1-p)^{n-t} \cdot \frac{\sum_{i=t-r-n_0}^{\min(k-n_0,t-n_0)}(i+2n_0)\binom{k-n_0}{i}\binom{r}{t-i-n_0}}{\binom{n-n_0}{t-n_0}}$$
$$= \frac{1}{2k}\sum_{t=r+n_0+1}^{n}\frac{\prod_{i=n-n_0+1}^{n} i}{\prod_{j=t-n_0+1}^{t} j}p^t(1-p)^{n-t} \cdot \sum_{i=t-r-n_0}^{\min(k-n_0,t-n_0)}(i+2n_0)\binom{k-n_0}{i}\binom{r}{t-i-n_0}.$$

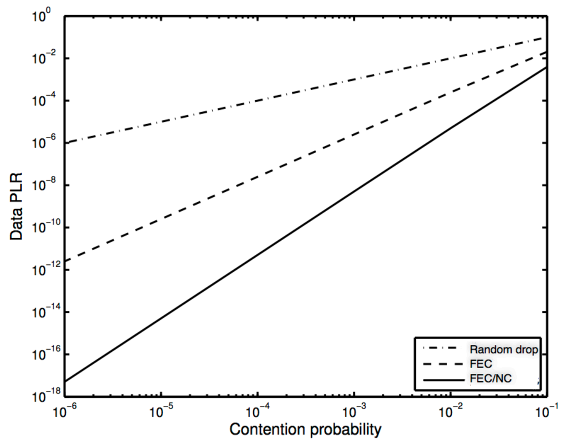
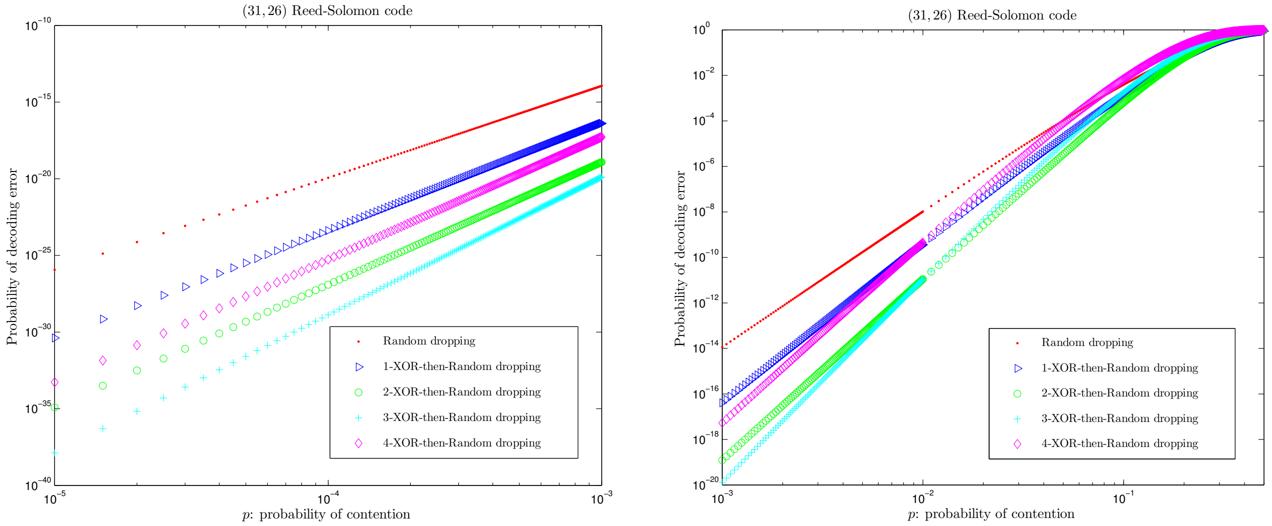


Fig. 3. Packet loss rates for RS(8,7)

C. Realistic RS code parameters and packet loss rates

The RS code rate for data transmission applications are generally set to $1/2$, if there is no information about the channel's loss characteristics. This is well in line with the finding of Theorem 1, which actually states that this is a minimum requirement for r for our scheme to be beneficial. However, Theorem 1 and 2 only assume a contention probability $p < 1/2$. In the OPS case, most researchers mention realistic contention probability values in the range of $10^{-5} < p < 10^{-2}$. In this case, an RS code with a higher code rate suffices, e.g., RS(8,7) already unlocks the potential of NC over random dropping. The exact values for n and k (and thus r) are also dictated by the finite buffering capabilities of edge routers (for decoding), as well as the incurred delay and delay requirements. In general, the higher the code rate, the better the bandwidth utilization; also, the shorter the code block n, the lower the incurred delay for decoding. It is also within the realm of possibilities to assign different (n, k) code parameters to different traffic classes.

Packet loss rates under a realistic range of contention probabilities are shown in Fig. 3; we use RS(8,7) in this example. Default OPS operation (without FEC or network coding) is plotted as a baseline. Note the significant improvement from the baseline through FEC to FEC/NC. In the case of $p = 10^{-2}$, for example: the baseline data packet loss rate is $\text{PLR} = 10^{-2}$,



(a) low contention probability p

(b) high contention probability p

Fig. 4. Decoding error probabilities for RS(31,26)

FEC can improve this value to $\mathrm{PLR}_{\mathrm{FEC}} = 2.45 \cdot 10^{-4}$; while, FEC/NC further reduces it to $\mathrm{PLR}_{\mathrm{NC}} = 4.92 \cdot 10^{-6}$. As it can be seen in the log-log plot, this trend is consistent across different contention probabilities, as the proposed combined FEC/NC scheme reduces the packet loss rate with several orders of magnitude. Note that with such small contention probabilities and a relatively short code block n, having more than 2 contentions inside a packet set has a negligible probability, hence we exclude XORing more than once.

We also investigate the impact of choosing n_0 on the probability of unsuccessful decoding; when the first n_0 contentions of DPs are XORed. For this purpose, we use RS(31,26) to have a longer code block allowing for multiple contentions inside a packet set. Results are plotted in Fig. 4, where we have separated the low (Fig. 4(a)) and high (Fig. 4(b)) contention probability cases for visual clarity. When p is low, 3-XOR-then-random dropping is the best option for the given code parameters. This means that there are not enough contentions inside a packet set to use 4-XOR optimally; on the other hand, 2-XOR cannot unleash the full potential of redundancy transfer. As we move to the higher contention region (around $p = 10^{-2}$), 2-XOR becomes optimal in terms of providing the lowest probability of unsuccessful decoding.

IV. Discussion

We have presented a first step towards applying combined FEC and NC in a bufferless network environment, specifically in OPS; naturally, our work have some limitations. First, we do not deal with the actual physical implementation of the proposed scheme. Whether we can use all-optical XOR gates [18] and optical signal processing [19] needs to be investigated taking into account coherent and non-coherent optical transmission modes. Second, we have only considered and analyzed the case of a single optical packet switch. To be practical in most scenarios, our scheme has to be extended to cover multiple switches and the signaling between them; potentially utilizing an SDN-style control plane. Third, our scheme targets flows with the same destination within the OPS network. However, OPS data center networks can be configured, virtual machines can be deployed and traffic can be steered accordingly, maximizing the gain of the coding mechanism.

As for future work in the OPS domain, studying networks with relevant real-world data center and core network topologies, a large number of flows and no flow synchronization using simulation, and designing a coordinated coding mechanism for multiple packet switches are all important next steps. Furthermore, performance and cost comparison with the numerous existing contention resolution schemes should also be addressed. Another line of future work could look into other potential application domains such as on-chip and industrial sensor-actuator networks.

V. Conclusion

In this paper we have proposed a combined forward error correction and network coding scheme to mitigate performance issues arising in bufferless networks. Specifically, we have presented a case study for alleviating packet loss in optical packet switched networks. Our analysis has shown that the combined coding scheme indeed has potential in increasing successful decoding probability and reducing data packet loss rate with orders of magnitude, if the underlying FEC exhibits sufficient redundancy. Factoring in realistic contention probabilities in OPS networks, high rate, low overhead RS codes can be utilized which are able to both reduce packet loss and meet buffering and decoding delay requirements at the same time. We believe that such a combined coding scheme has the potential to be utilized in OPS (data center and core networks) and other domains such as on-chip and industrial networks, where (near-)zero buffers are required.

ACKNOWLEDGEMENTS

Gergely Biczók has been supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

REFERENCES

[1] K. Koo, T. Ha, N. An, K.-J. Park, and H. Lim, “Zero-buffer data delivery for industrial networks,” in *Ubiquitous and Future Networks (ICUFN), 2015 Seventh International Conference on*. IEEE, 2015, pp. 100–102.

[2] T. Moscibroda and O. Mutlu, “A case for bufferless routing in on-chip networks,” in *ACM SIGARCH Computer Architecture News*, vol. 37, no. 3. ACM, 2009, pp. 196–207.

[3] C.-L. Chou and R. Marculescu, “Contention-aware application mapping for network-on-chip communication architectures,” in *Computer Design, 2008. ICCD 2008. IEEE International Conference on*. IEEE, 2008, pp. 164–169.

[4] G. Michelogiannakis, D. Sanchez, W. J. Dally, and C. Kozyrakis, “Evaluating bufferless flow control for on-chip networks,” in *Proceedings of the 2010 Fourth ACM/IEEE International Symposium on Networks-on-Chip*. IEEE Computer Society, 2010, pp. 9–16.

[5] D. K. Hunter and I. Andonovic, “Approaches to optical internet packet switching,” *Communications Magazine, IEEE*, vol. 38, no. 9, pp. 116–122, 2000.

[6] S. Yao, B. Mukherjee, S. B. Yoo, and S. Dixit, “A unified study of contention-resolution schemes in optical packet-switched networks,” *Lightwave Technology, Journal of*, vol. 21, no. 3, pp. 672–683, 2003.

[7] S. B. Yoo, “Energy efficiency in the future internet: the role of optical packet switching and optical-label switching,” *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 17, no. 2, pp. 406–418, 2011.

[8] J. Perelló, S. Spadaro, S. Ricciardi, D. Careglio, S. Peng, R. Nejabati, G. Zervas, D. Simeonidou, A. Predieri, M. Biancani *et al.*, “All-optical packet/circuit switching-based data center network for enhanced scalability, latency, and throughput,” *Network, IEEE*, vol. 27, no. 6, pp. 14–22, 2013.

[9] W. Miao, F. Agraz, S. Peng, S. Spadaro, G. Bernini, J. Perelló, G. Zervas, R. Nejabati, N. Ciulli, D. Simeonidou *et al.*, “Sdn-enabled ops with qos guarantee for reconfigurable virtual data center networks,” *Journal of Optical Communications and Networking*, vol. 7, no. 7, pp. 634–643, 2015.

[10] B. Guo, S. Peng, C. Jackson, Y. Yan, Y. Shu, W. Miao, H. J. Dorren, N. Calabretta, F. Agraz, S. Spadaro *et al.*, “Sdn-enabled programmable optical packet/circuit switched intra data centre network,” in *Optical Fiber Communication Conference*. Optical Society of America, 2015, pp. Th4G–5.

[11] H. Øverby, “Network layer packet redundancy in optical packet switched networks,” *Optics express*, vol. 12, no. 20, pp. 4881–4895, 2004.

[12] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, “Xors in the air: practical wireless network coding,” in *ACM SIGCOMM computer communication review*, vol. 36, no. 4. ACM, 2006, pp. 243–254.

[13] A. Zhan and C. He, “Joint design of channel coding and physical network coding for wireless networks,” in *Neural Networks and Signal Processing, 2008 International Conference on*. IEEE, 2008, pp. 512–516.

[14] F. Chang, K. Onohara, and T. Mizuochi, “Forward error correction for 100 g transport networks,” *Communications Magazine, IEEE*, vol. 48, no. 3, pp. S48–S55, 2010.

[15] E. D. Manley, J. S. Deogun, L. Xu, and D. R. Alexander, “All-optical network coding,” *Journal of Optical Communications and Networking*, vol. 2, no. 4, pp. 175–191, 2010.

[16] F. J. MacWilliams and N. J. A. Sloane, *The theory of error correcting codes*. Elsevier, 1977, vol. 16.

[17] J. Lacan, V. Roca, J. Peltotalo, and S. Peltotalo, “Reed-solomon forward error correction (fec) schemes, april 2009, ietf request for comments,” RFC 5510 (Standards Track/Proposed Standard), Tech. Rep.

[18] M. Zhang, L. Wang, and P. Ye, “All optical xor logic gates: technologies and experiment demonstrations,” *Communications Magazine, IEEE*, vol. 43, no. 5, pp. S19–S24, 2005.

[19] Z. Liu, M. Li, L. Lu, C.-K. Chan, S.-C. Liew, and L.-K. Chen, “Optical physical-layer network coding,” *IEEE Photonics Technology Letters*, vol. 24, no. 13, p. 1424, 2012.