# Parity-Based Loss Recovery for Reliable Multicast Transmission

J. Nonnenmacher, E Biersack, D. Towsley

# Objectives

- Investigate how forward error correction (FEC) can be **combined** with Automatic Repeat Request (ARQ) to achieve scalable reliable multicast.
- Two scheme: Layered Multicast and Integrated Multicast
- Consider different loss models (e.g., independent, spatially or temporally correlated, homogeneous and heterogeneous loss receivers).

# Introduction

- Consider video/audio distributions and collaborative works (e.g., vic, vat, wb under MBONE).
- Many issues, one important one is how to provide a *scalable* and *reliable* multicast.
- To achieve reliable multicast, two techniques:
  - ARQ: requests for retransmission of lost packets. Watch out for feedback explosion problem).
  - FEC: transmits redundant **parity** data along with original data. If the amount of original data lost is not more than the amount of parity, we can perform perfect reconstruction.

- Previous literature on reliable multicast considered techniques like slotting/damping and hierarchy to avoid feedback explosion. Both of these approaches need parameters tuning and careful multicast design.
- Other considered FEC but the *scalability* was small.

# 2: How FEC works

General idea:

- Given $k$ data packets $d_1, \ldots, d_k$, we use techniques such as Reed Solomon Erasure code to generate parity packets $p_1, \ldots, p_h$, where $h = n - k$.
- receiver can reconstruct the data packets $d_1, \ldots, d_k$ whenever it receives any $k$ packets from $d_1, \ldots, d_k, p_1, \ldots, p_h$.
- The $k$ data packets forms a transmission group (**TG**).
- The $n$ packets forms an **FEC block**.

- In practice, we send $k$ data packets first and then the parity packets.
- This simplifies decoding because:
  - If a receiver gets $k$ data packets correctly. This implies no decoding overhead.
  - if $l < n - k$ out of the $k$ data packets are lost, the decoding overhead is *proportional* to $l$.

Benefits of using parity packets for multicast transmission:

- a single parity can be used to repair the loss of any one of the $k$ data packets. It implies repairing *different* data packets at *different* receivers.
- Improved scalability in terms of group size. The reason is that feedback is reduced from per-packet feedback (to indicate which packet is lost) to per-TG feedback.

## 3: Placement of FEC in a R-Multicast

- Two approaches: Layered FEC and Integrated FEC.
- Layered FEC: add a layer responsible for FEC between the network layer and the reliable multicast layer.
- Integrated FEC: integrate FEC with reliable multicast in one layer

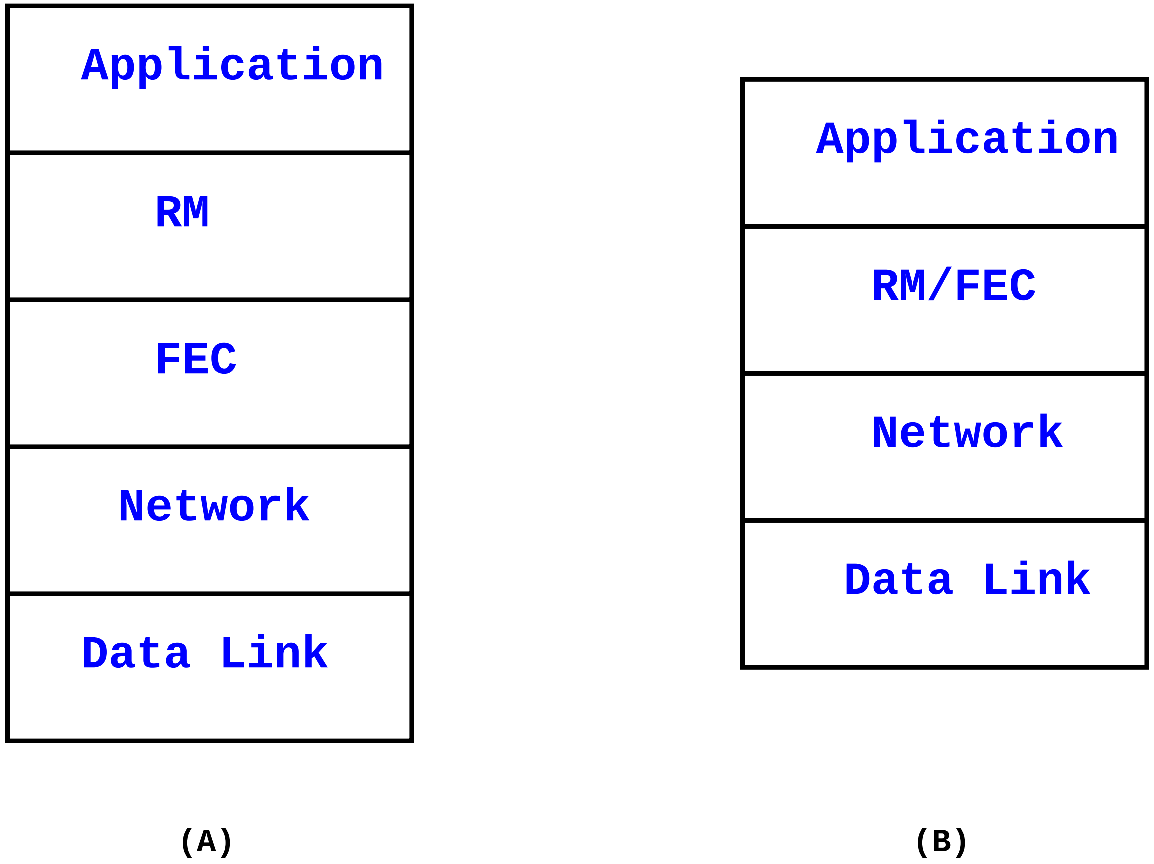


Figure 1: (A) Layered FEC; (B) Integrated FEC

Advantages of layered FEC multicast:

- separation of FEC and reliable multicast
- designer can focus on different components
- FEC layer can be used for other applications

Is there any potential advantage of the Integrated FEC? We want to evaluate the performance of (a) no FEC, (b) layered FEC, (c) integrated FEC. The performance measure is the **average number of transmissions required to transmit a packet reliably to all receivers**. This measure relates

- the bandwidth overhead
- processing overhead

# Layered FEC: independent loss

Packet losses occur as independent events, both spatially and temporally with probability $p$.

- RM passes the data packets to the FEC layer.
- FEC layer takes $k$ data packets, produces $h$ parities packets, sends all these $n$ $(= k + h)$ packets.
- Upon receiving, receives at least $k$ out of $n$ packets, delivered $k$ data packets to the RM layer.
- If less than $k$, reconstruction is not possible. FEC discards parity packets. The sending RM retransmits the lost original as part of a new group.

- Let $q(k,n,p)$ be the probability that the RM receiver does not receive a random data packet from the TG (transmission group). We have

$$q(k,n,p) = p\left[1 - \sum_{j=0}^{n-k-1} \left( \begin{array}{c} n-1 \\ j \end{array} \right) p^j (1-p)^{n-j-1}\right] \quad 1 \le k \le n.$$

- Let $M'$ denote the random variable of number of times an arbitrary data packet is transmitted before all RM receivers have received it. We have:

$$P(M' \le i) = \left[1 - q(k,n,p)^i\right]^R \qquad i \ge 1.$$

- Let $M$ be the random variable equivalent to $M'$ with the additional accounting of the parity packets added at the FEC layer. We have:

$$
\begin{aligned}
E[M] &= \frac{n}{k} E[M'] \\
&= \frac{n}{k} \sum_{i=0}^{\infty} 1 - P(M' \leq i)
\end{aligned}
$$

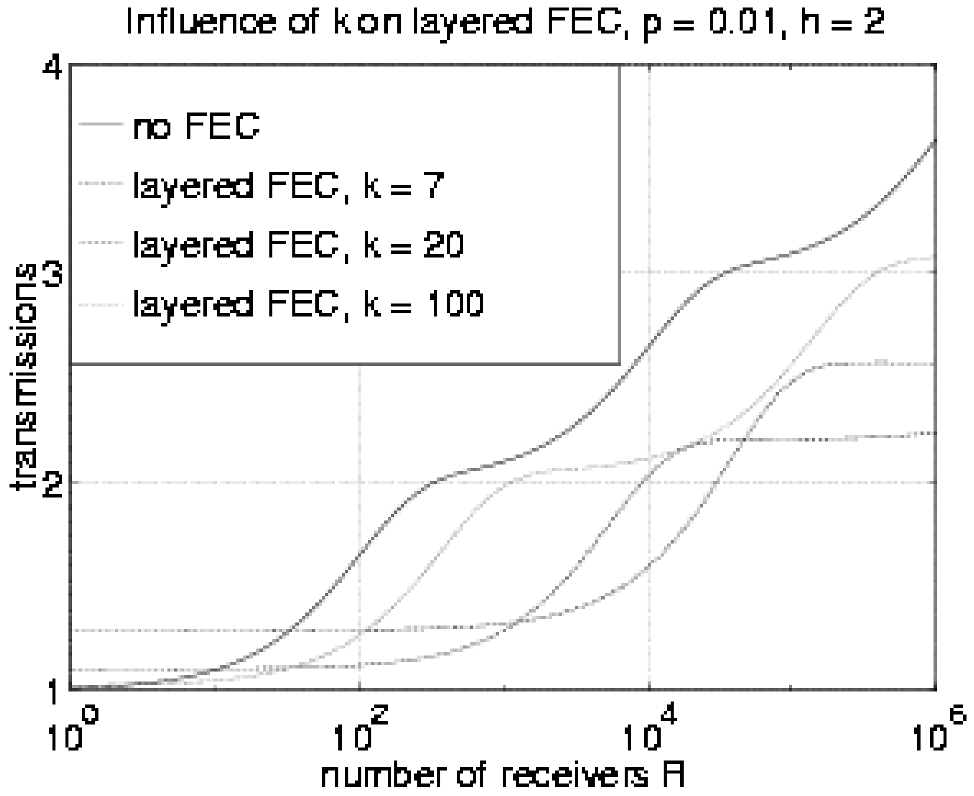


**Figure 3: Non FEC versus layered FEC with $h = 2$ parity packets for different TG sizes $k = 7, 20, 100$ and loss probability $p = 10^{-2}$.**

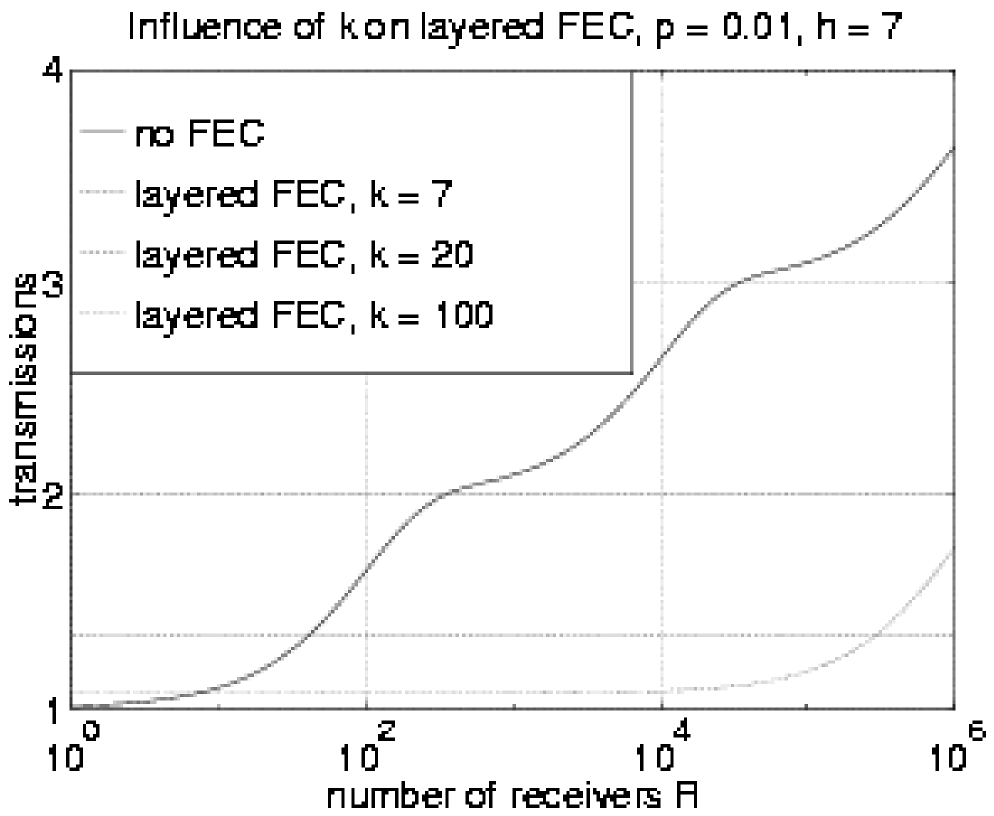


**Figure 4: Non FEC versus layered FEC with $h = 7$ parity packets for different TG sizes $k = 7, 20, 100$ and loss probability $p = 10^{-2}$.**

# Observation of Layered FEC

- a decrease in the expected number of transmissions when FEC is introduced and the receiver population is large.
- the number of parity packets needs to be matched to the transmission group. E.g., when $k = 100$ with $h = 2$, the performance is worse than when TGs of sizes 7 or 20 with the same parity packet. On the other hand, when TG of size 100 coupled with $h = 7$ performs better than TGs of size 7 and 20 for receiver populations in 1-200,000 range.

# Integrated FEC

Many ways to implement Integrated FEC, we consider:

- the sender sends a TG along with $a \leq h$ parity packets.
- Each time a receiver detects a missing packets, sender multicasts a new parity packet until receiver has sufficient number of packets (e.g., $k$ out of $n$).
- If all parity packets are used up, packets requiring retransmission are placed into a new TG.

**Case 1:**

We first derive an unachievable lower bound of $E[M]$. This corresponds when $n = \infty$. Let $L_r$ denote the number of *additional* packet transmissions required by a random receiver. We have

$$
\begin{aligned}
P(L_r = 0) &= \sum_{j=0}^{a} \binom{k+a}{j} p^j (1-p)^{k+a-j} \\
P(L_r = m) &= \binom{k+a+m-1}{k-1} p^{m+a} (1-p)^k \qquad m = 1, 2, ..
\end{aligned}
$$

Let $L$ be the maximum number of additional packets that are transmitted. We have

$$P(L \leq m) = [P(L_r \leq m)]^R = \left[\sum_{i=0}^{m} P(L_r = i)\right]^R \quad m = 1, 2, ..$$

The mean number $E[L]$ is:

$$E[L] = \sum_{m=0}^{\infty} (1 - P(L \leq m))$$

Finally, the average number of transmission is:

$$E[M] = \frac{E[L] + k + a}{k}$$

**Case 2:**

When $n < \infty$, let $B$ denote the number of times that an arbitrary packet was transmitted (i.e., the number of blocks transmitted that included it). Note that the transmission of the first $(B-1)$ groups require the transmission of $n$ packets each. For the last group, the number of packets is a r.v. of $L$ given that $L \leq n$. Therefore, we have

$$
\begin{aligned}
E[M] \quad = \quad & \frac{n}{k}\left(\sum_{i=0}^{\infty}(1-(1-q(k,n,p)^{i})^{R}-1\right) + \\
& \frac{1}{k}\left(\sum_{m=0}^{n}(1-P(L \leq m|L \leq n)) + k + a\right)
\end{aligned}
$$

(Note: it is different from the paper.)

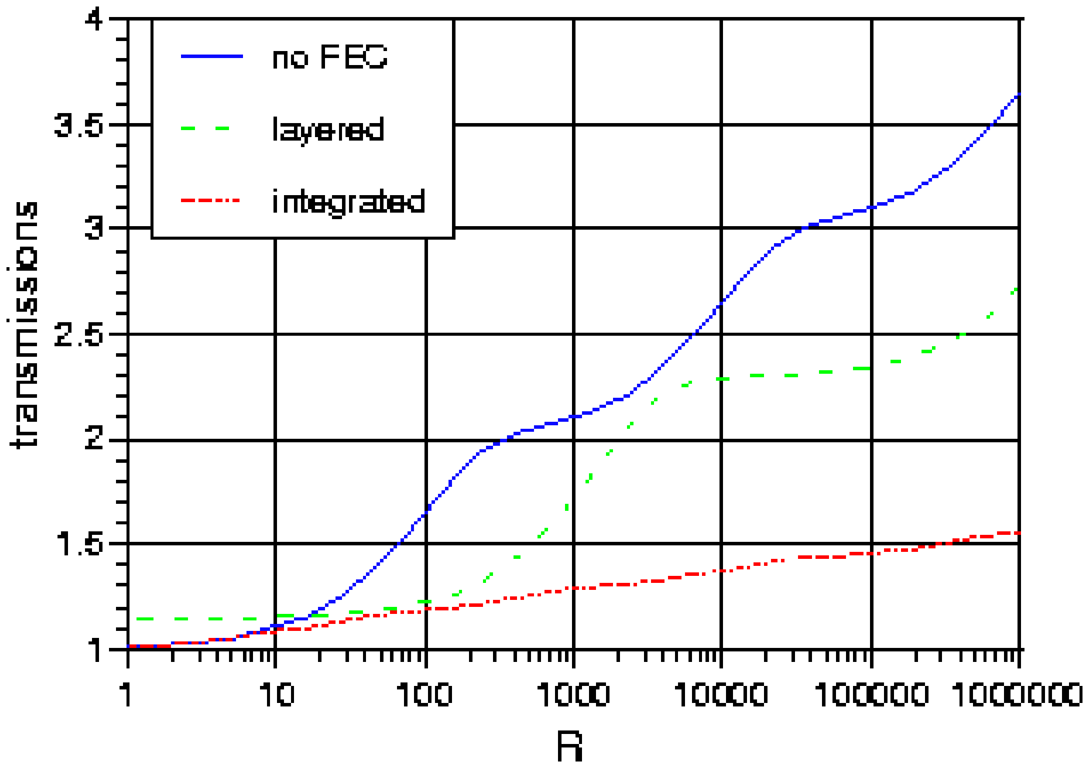


**Figure 5:** $E[M]$ **versus** $R$ **for TG size of 7 and loss probability of** $p = 10^{-2}$**. Layered FEC versus integrated FEC.**

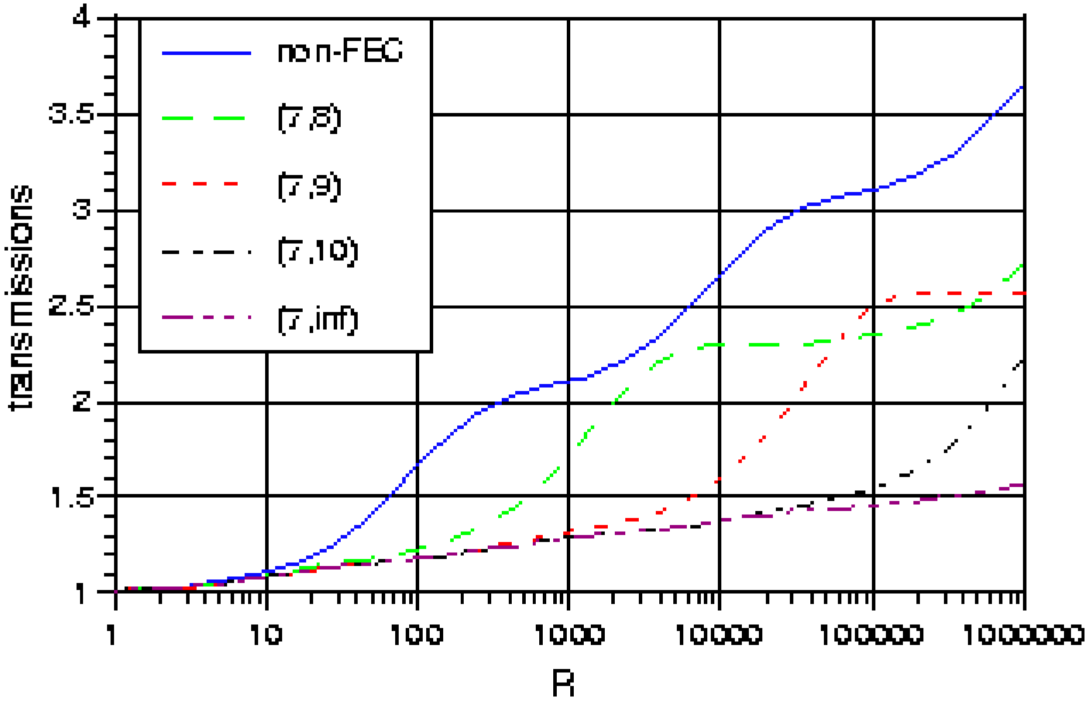


**Figure 6:** $E[M]$ versus $R$ for TG size of 7 and loss probability of $p = 10^{-2}$. Integrated FEC with $k = 7$ for different values of $h$.

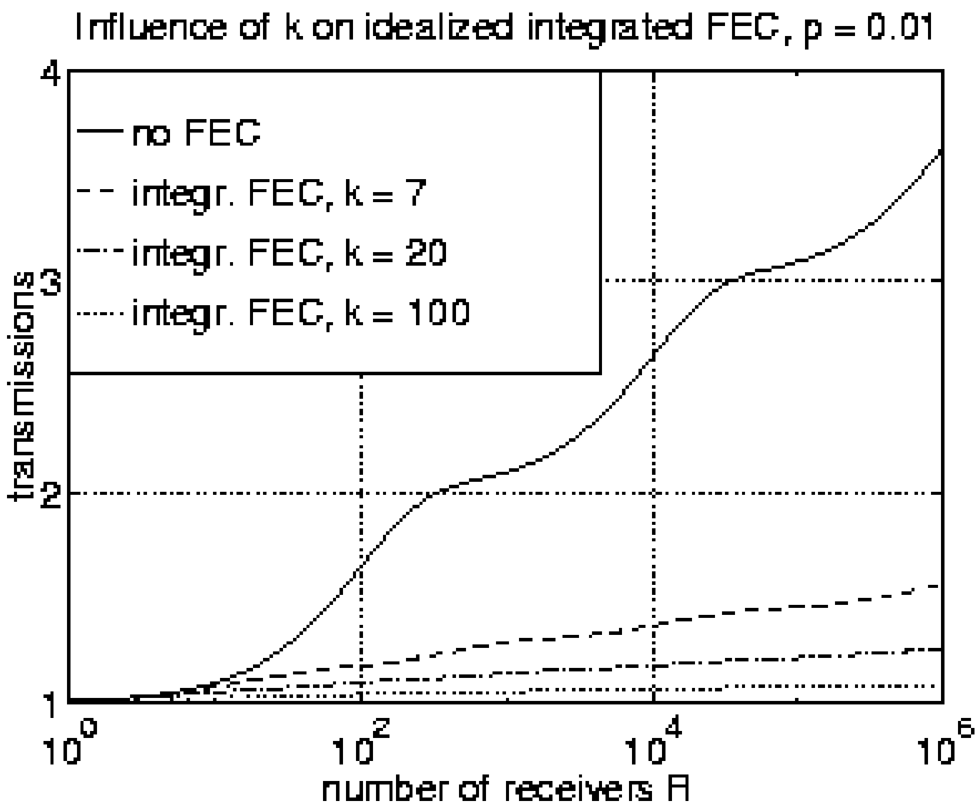


**Figure 7: Influence of number of receivers on integrated FEC for different TG sizes for packet loss probability $p = 10^{-2}$.**

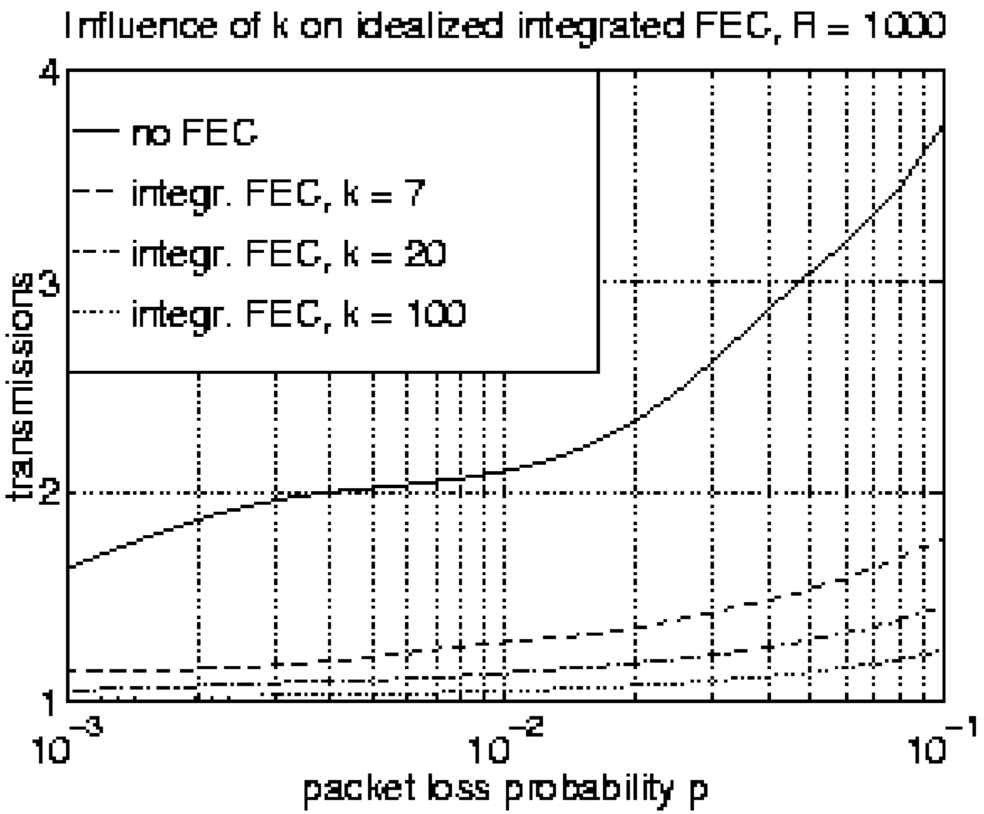


**Figure 8: Influence of loss probability on integrated FEC for different TG sizes for $R = 1000$ receivers.**

## Results:

- Figure 5 illustrates that the integrated FEC has the potential for a large performance improvement over layered FEC.
- Figure 6 illustrates how many parity packets are needed to have a performance close to the lower bound. It shows that when $k = 7, h = 3$, we are close to the performance lower bound with large scalability support. Hence, authors just use lower bound for comparison.
- Figure 7 and 8 show the *insensitivity* to the number of receivers $R$ and loss probability $p$. Increasing the TG size has little effect on the performance.

## Heterogeneous Receivers

Consider receivers with different lost probabilities. Let $p(r)$ be the lost probability of receiver $r = 1, \ldots, R$. Assume losses are spatially and temporally independent. For **layered FEC**

$$E[M] = \frac{n}{k} \sum_{i=0}^{\infty} \left( 1 - \prod_{r=1}^{R} \left( 1 - q(k,n,p(r))^{i} \right) \right)$$

For **Integrated FEC**, it is similar except using individual loss probabilities.

Let us consider a population of two classes of receivers, $R(1-\nu)$ with $p(r) = 10^{-2}$ and $R\nu$ with $p(r) = 0.25$. Authors vary $\nu$ to see the effect.

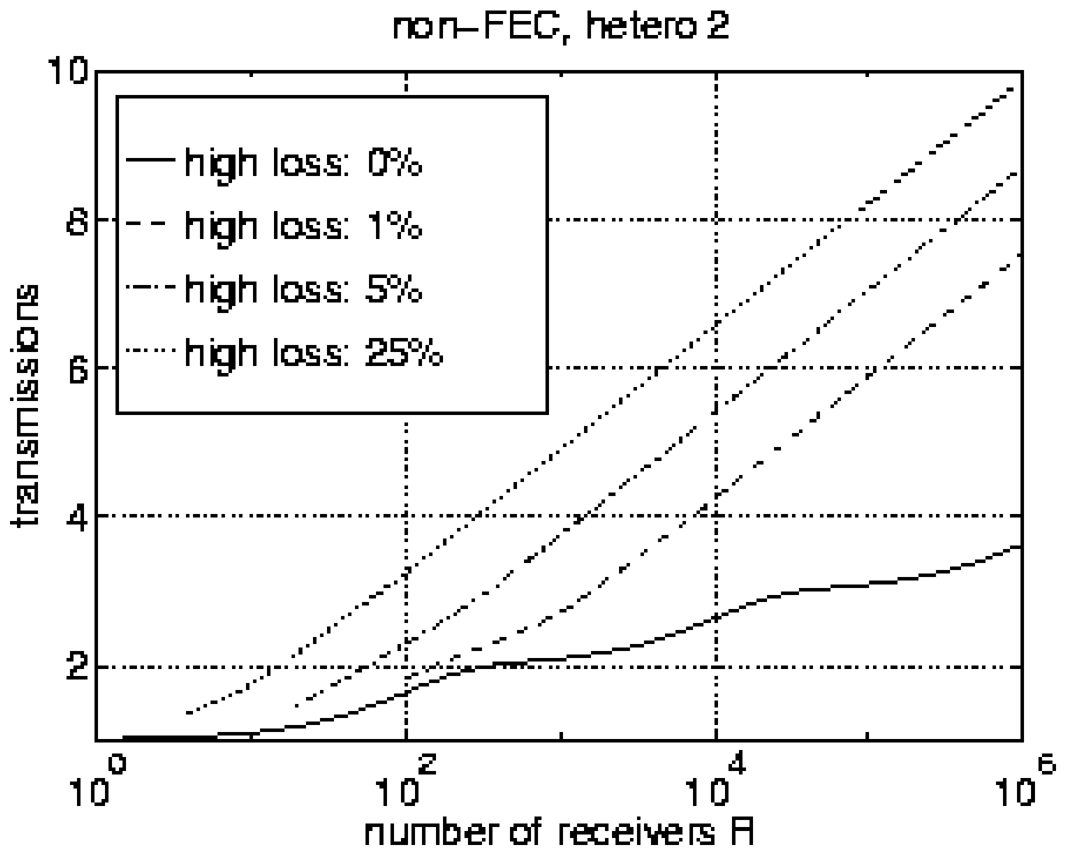


**Figure 9: Reliable multicast for different heterogenities without FEC.**

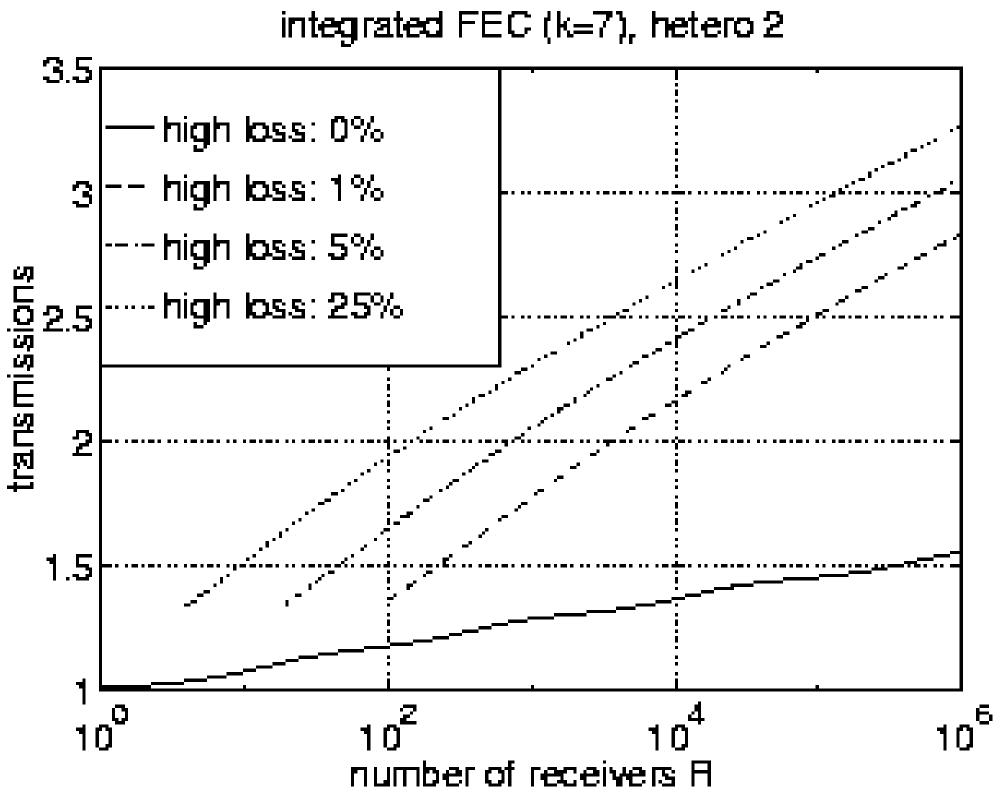


**Figure 10: Reliable multicast for different heterogenities with integrated FEC.**

# Result

- influence of high loss receivers increases with the number of receivers.
- For a group of 1 million receivers, the presence of 10,000 high loss receivers (1%) is sufficient to double the number of transmissions.
- For small group (e.g., 100 receivers), high loss receivers have a much less effect on the expected number of transmissions.

In a real multicast group, high loss receivers is in most cases determined by the position of a high loss router in the tree. Let us look at correlated loss.

# Spatially Correlated Loss

Two loss models:

- **Independent loss:** only receivers lose packets, each receiver independently with probability $p$. Other nodes in the multicast tree do not lose packets.
- **Full binary tree shared loss:** A full binary tree with height $d$, where losses occur as independent event at each node with probability $p_{node}$. The source is the root and receivers are the leaves. $p_{node}$ is chosen such that:

$$p = 1 - (1 - p_{node})^{d+1}$$

Note each receiver experiences the *same* loss probability.

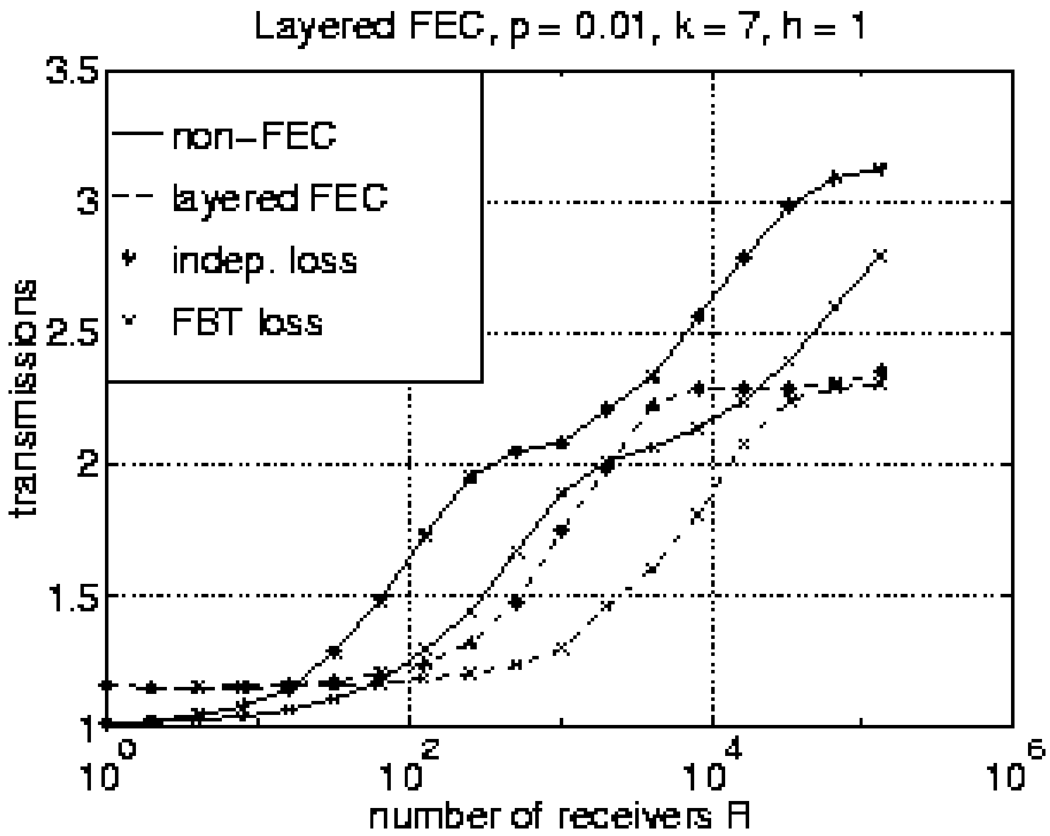


**Figure 11: Layered FEC compared with non-FEC for Independent and for FBT Shared Loss.**

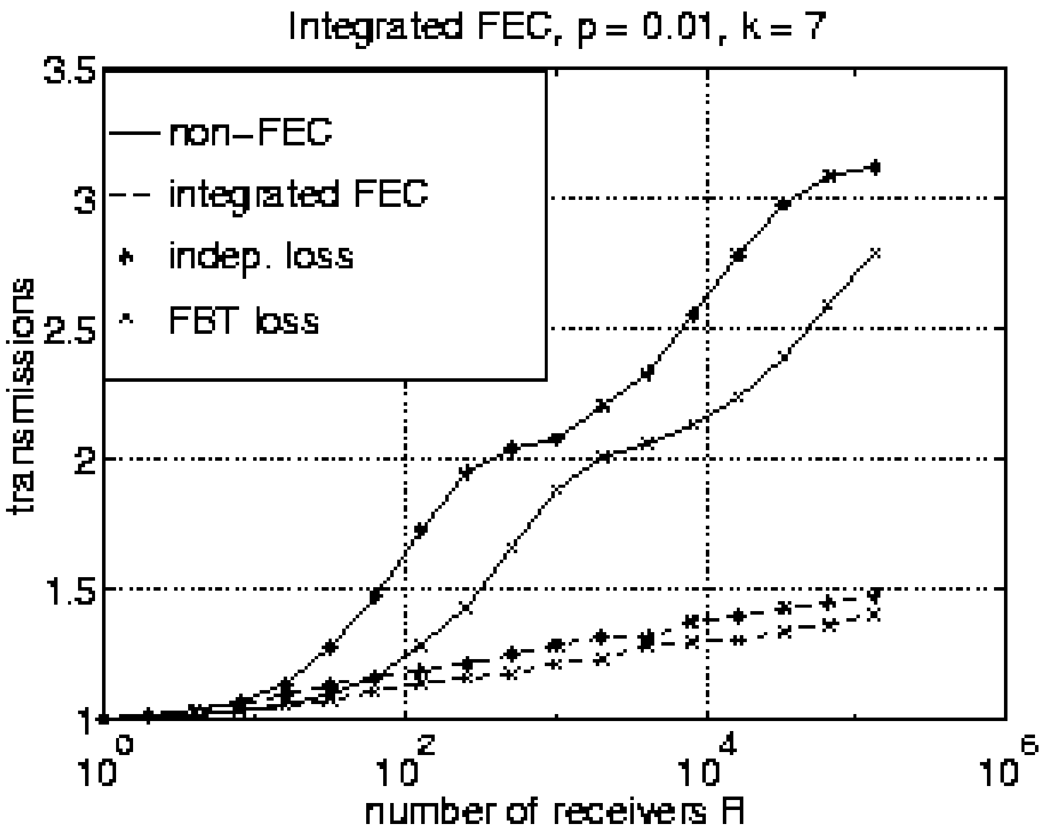


**Figure 12: Integrated FEC compared with non-FEC for Independent and for FBT Shared Loss.**

## Result

- Layered FEC transmits parties whether or not a loss occurs When shared loss occurs, the parity does not exhibit the same repair efficiency.
- Figure 11 shows that for independent loss, layered FEC amortized for $R > 20$. But for shared loss, it amortized at $R > 60$. It means FEC is only good for a multicast group of certain size.
- The number of transmissions is lower (often substantially) when losses are shared than independent.
- Figure 12 shows that integrated FEC is insensitive to shared or independent losses.

# Temporal Loss

Packet losses are described by a two state continuous time Markov chain $\{X_t\}$ where $X_t \in \{0, 1\}$. Packet transferred at time $t$ is lost if $X_t = 1$ and not lost if otherwise.

Note that the timing of the retransmissions has an influence on the performance of the loss recovery. Author considered the following:

- **No FEC:** the time between successive transmissions of the same packet to be spaced by time $\Delta + T$.
- **Layered FEC:** the time between the sending of the last packet of a FEC block and the time of the sending of the first packet in the successive FEC block, containing the retransmission, to be spaced by time $\Delta + T$.

- **Integrated FEC 1:** Parity packets are transmitted with the same rate $1/\Delta$ immediately following the original packets. When a receiver has received enough parity packets, it leaves the multicast group. In this scheme, no feedback is needed for loss recovery and there is no unnecessary delivery and reception of parity packets, provided that the time needed to depart from the group is smaller than the packet inter-arrival time.
- **Integrated FEC 2:** receivers send NAKs indicating the number of missing packets. Feedback is sent after the transmission of the original packets, after the first retransmission of parities, etc. Subsequently the sender transmits the maximum number of parity packets needed by any receiver.

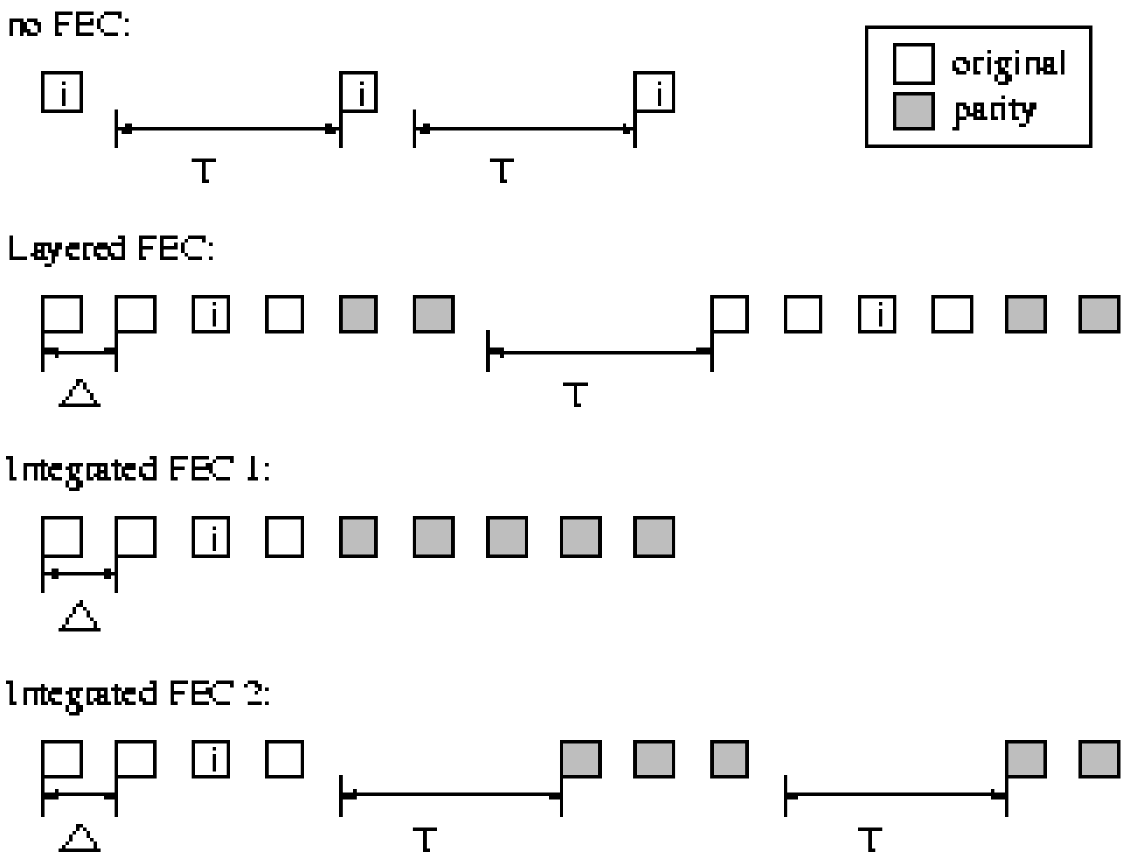


**Figure 13: Timing considerations of the different approaches.**

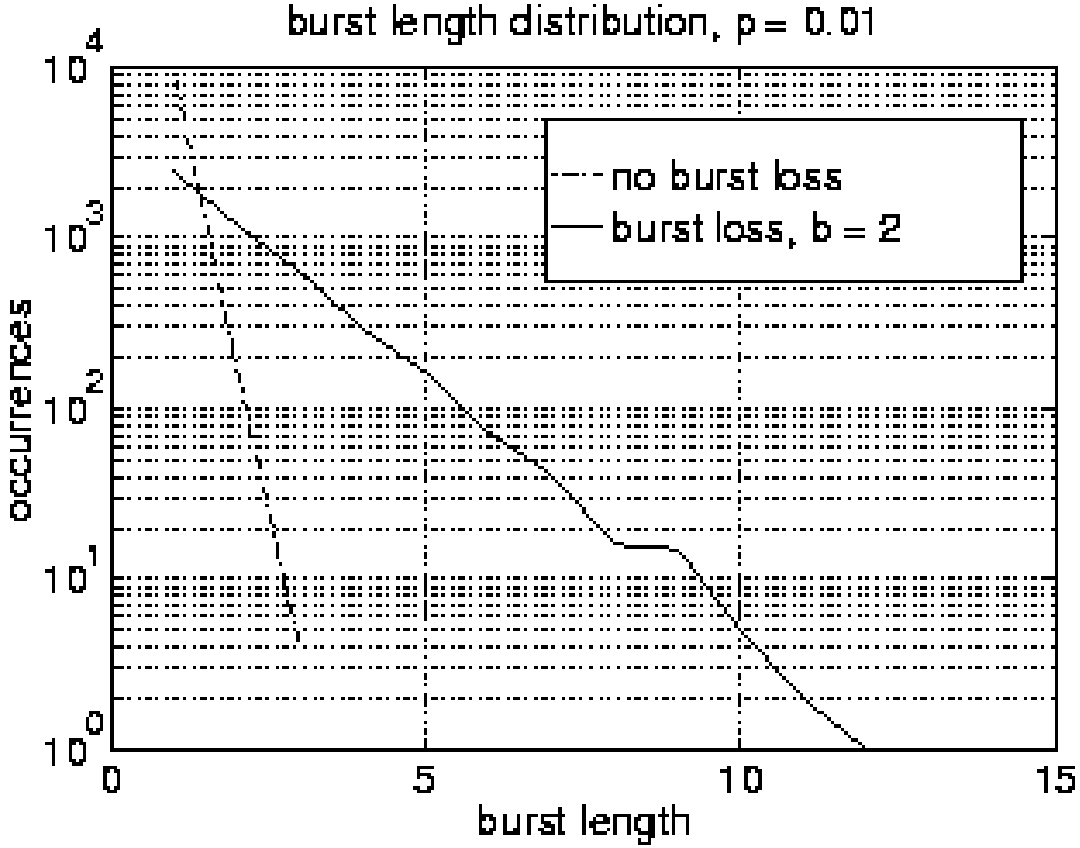


**Figure 14: Distribution of number of consecutive losses at one receiver, for no burst and burst ($\bar{b} = 2$) for a packet loss probability of $p = 0.01$.**

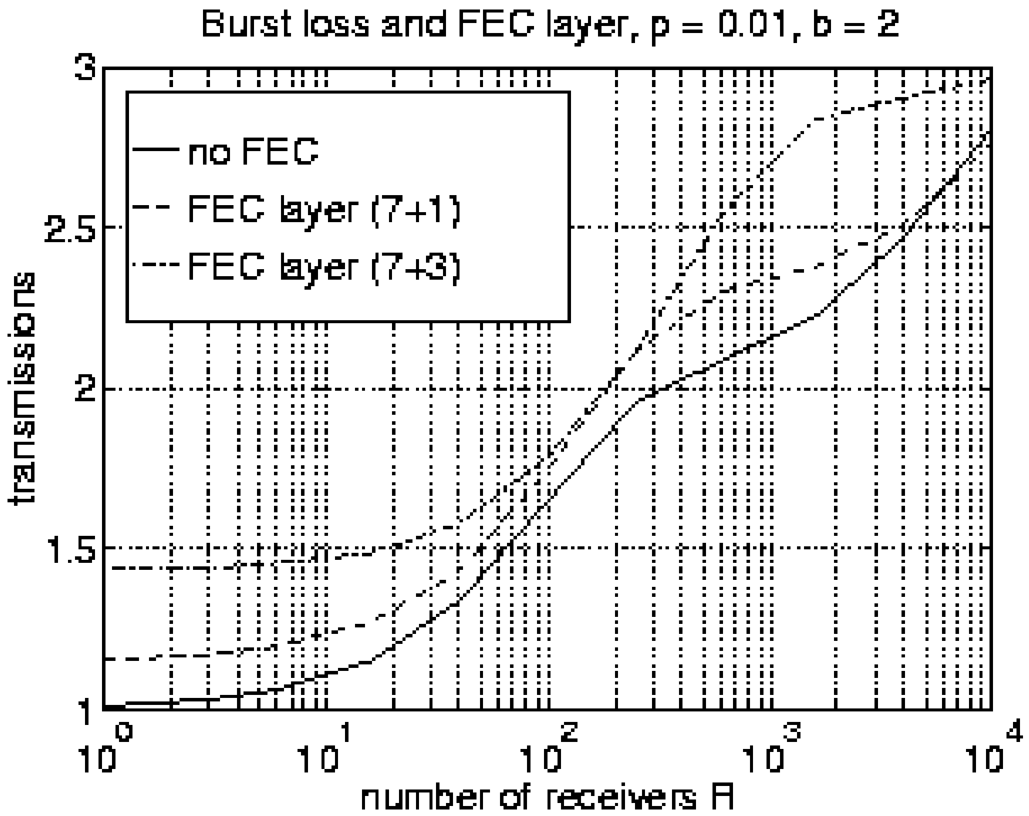


**Figure 15: Comparison of reliable multicast without FEC and with layered FEC for low ($h = 1$) and high ($h = 3$) redundancy for $k = 7$, $p = 0.01$ and $\bar{b} = 2$.**

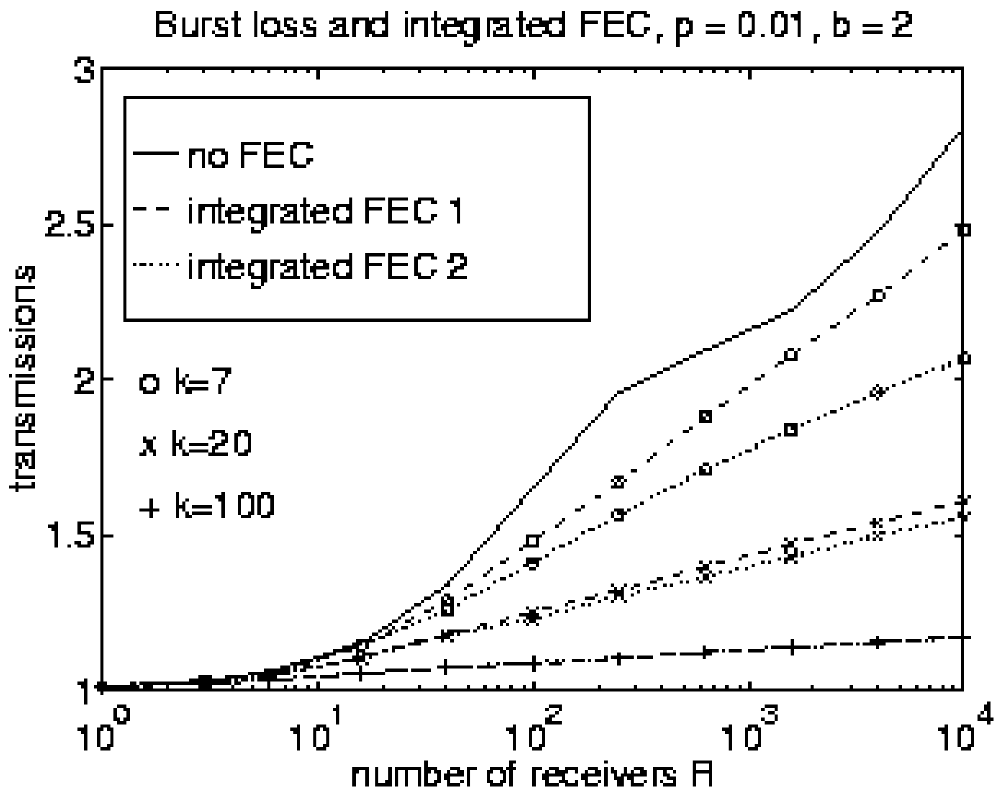


**Figure 16: Comparison of integrated FEC 1 and integrated FEC 2 for transmission group sizes $k = 7, 20, 100$, $p = 0.01$ and $\bar{b} = 2$.**

## Result

- Figure 14 shows the burst length density functions.
- Figure 15 shows that layered FEC actually may perform worse than no FEC.
- Integrated FEC still can perform better than no FEC under correlated losses. Also, we can increase the TG size to improve the performance.

# Summary

Using FEC in an integrated fashion provides five major benefits:

- it shifts resource usage from the network to the end-systems. It results in lower bandwidth at the cost of the end-system.
- The achievable end-system throughput for protocols based on Integrated FEC is higher than for non-FEC protocols, when data is pre-encoded.
- The error-control feedback is reduced.
- A moderate transmission group size of $k = 20$ will tolerate burst losses, even without interleaving.
- Good scalability.

# Conclusions

- Integrated FEC dramatically reduces the mean number of transmissions as compared to non-FEC.
- Integrated FEC is better than layered FEC for all parameters.
- Layered FEC can reduce the number of transmission needed for large receiver populations. However, its performance is sensitive to the coding parameters and the presences/absence of burst losses.
- High loss receivers determine the performance, even if they are a small subset of all receivers.
- With shared loss the whole repair efficiency of FEC is not used in the same extent than in the case of independent loss.
- For burst loss layered FEC is worse than no FEC.