MTTFPA considerations for 40GE and 100GE

Mark Gustlin and Mark Nowell

Cisco

IEEE HSSG November 2007

Topics

- Introduction
- Scrambling effects
- Burst Errors
- Summary

Overview of MTTFPA

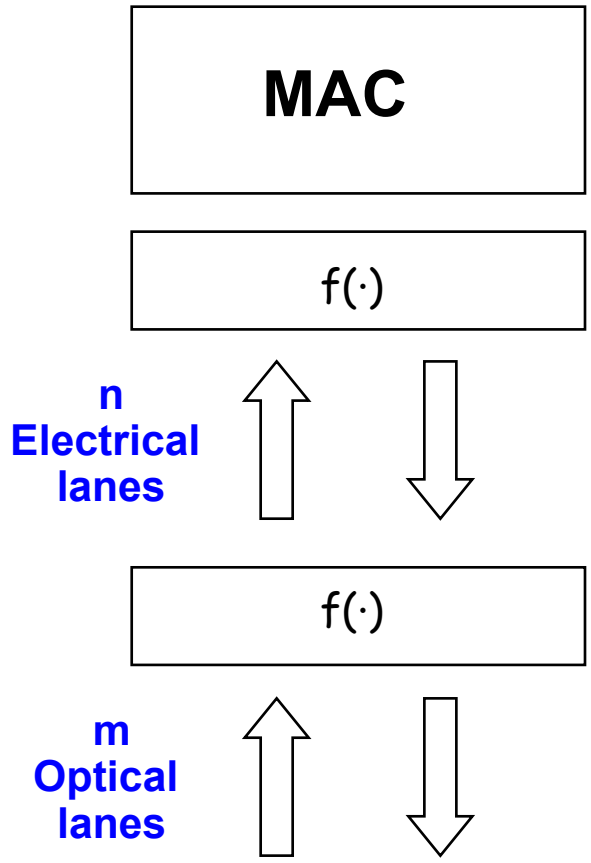
- **Mean Time To False Packet Acceptance is a figure of merit that needs to be understood for any proposed Ethernet architecture**
- **Typically, acceptable results are measured in billions of years**
- **Traditionally, Ethernet's CRC32 has inherent properties that provide acceptable MTTFPA**
- **Scrambling and striping both can have negative effects on CRC32's robustness and this must be analysed for 40GE and 100GE.**

MTTFPA and 40/100GE

- Ethernet's CRC32 has the following error detection capability
  - All 1, 2 or 3 bit errors are detectable
  - All burst up to 32 bits are detectable
  - All two burst errors up to 8 bits are detectable
- For the 10GBASE-R scrambler, single bit errors become 3 bit errors
- This was shown to not degrade the error detection capability of the IEEE CRC32 for 10GBASER (walker_1_0100)
  - No CRC degradation occurs if the CRC and the scrambler polynomial do not share common factors
  - The IEEE CRC32 has no common factors with the X^58 scrambler
  - Therefore, If the original errors can be detected, the multiplied errors can be detected
  - **The lack of common factor simplifies analysis**
- 40GE and 100GE requires striping of packet data, does that impact the MTTFPA?

Scrambling and Striping

Schematic example



Packets logically a single flow

Striping *may* occur here
Scrambling *may* occur here

Scrambling *may* occur here at a per lane level

Striping *may* occur here (if m≠n)
Scrambling *may* occur here

Scrambling *may* occur here at a per lane level

m does not have to equal n

Scrambling's Impact on MTTFPA

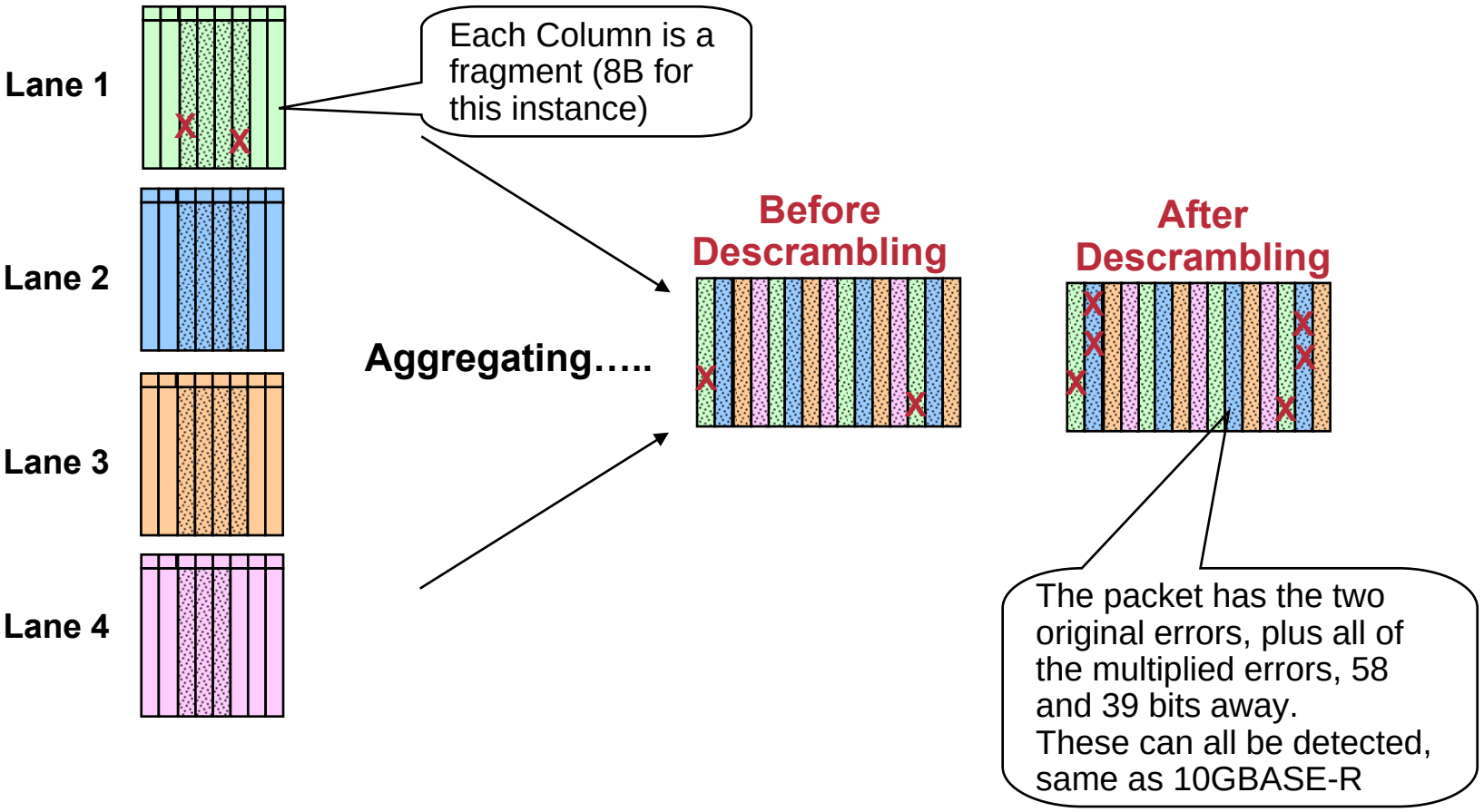
- Using a self synchronous scrambler multiplies errors
- How does the multiplication along with striping affect MTTFPA?
  - This depends on where scrambling (and therefore descrambling) is done
- For both 40GE and 100GE we will stripe packets across multiple lanes
  - Options discussed so far: 8B at a time (CTBI, PBL), 16B-256B (APL)
- We can scramble (and therefore descramble) the data in a number of places:
  1. At the 100G/40G aggregate level before distribution to multiple lanes
  2. On a per virtual lane (for CTBI)
  3. On a per electrical lane
  4. On a per optical lane

  Or a combination of the above

  Note that #2 and #3 are identical if the number of VLs = physical lanes

Scrambling at the Aggregate Level

- When data is scrambled and descrambled at the aggregate level (40G or 100G) then the MTTFPA analysis is simpler
- Example: Two random bit errors (8B data striping example, 4 lanes):



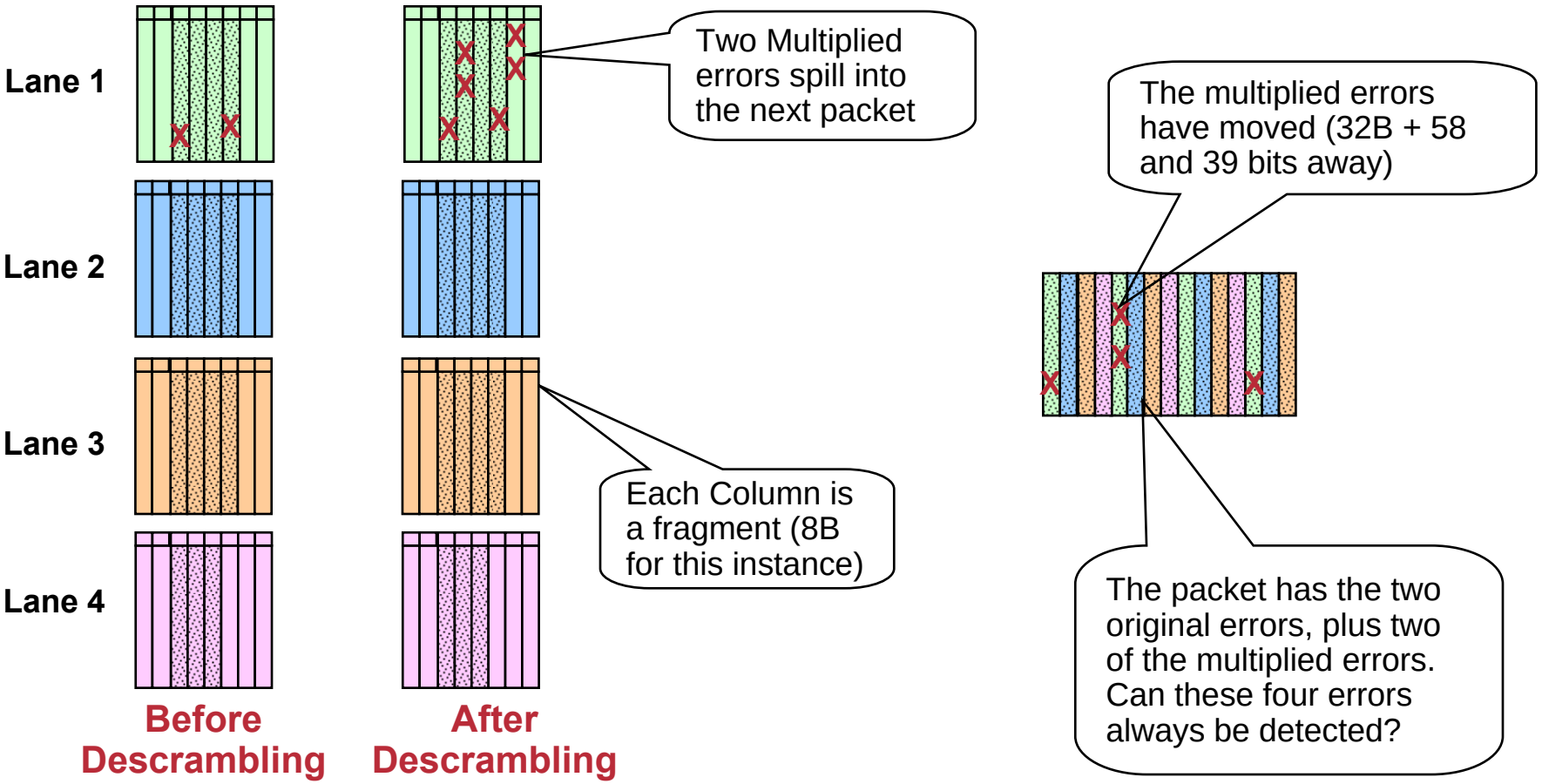
Scrambling at the Aggregate Level

- When data is scrambled and descrambled at the aggregate level (40G or 100G) then the MTTFPA analysis is similar to 10GBASE-R
- Looks just like the 10GBASE-R case, if the original error is detectable, then the multiplied errors are detectable
- Some boundary cases still have to be analyzed
- The above is true regardless of the fragmentation size, 8B, 16B etc…
- Errors are 100% detectable if the original errors are detectable, same as 10GBASE-R

Conclusion: Scrambling at aggregate level means the MTTFPA analysis is similar to that done for 10GBASE-R.

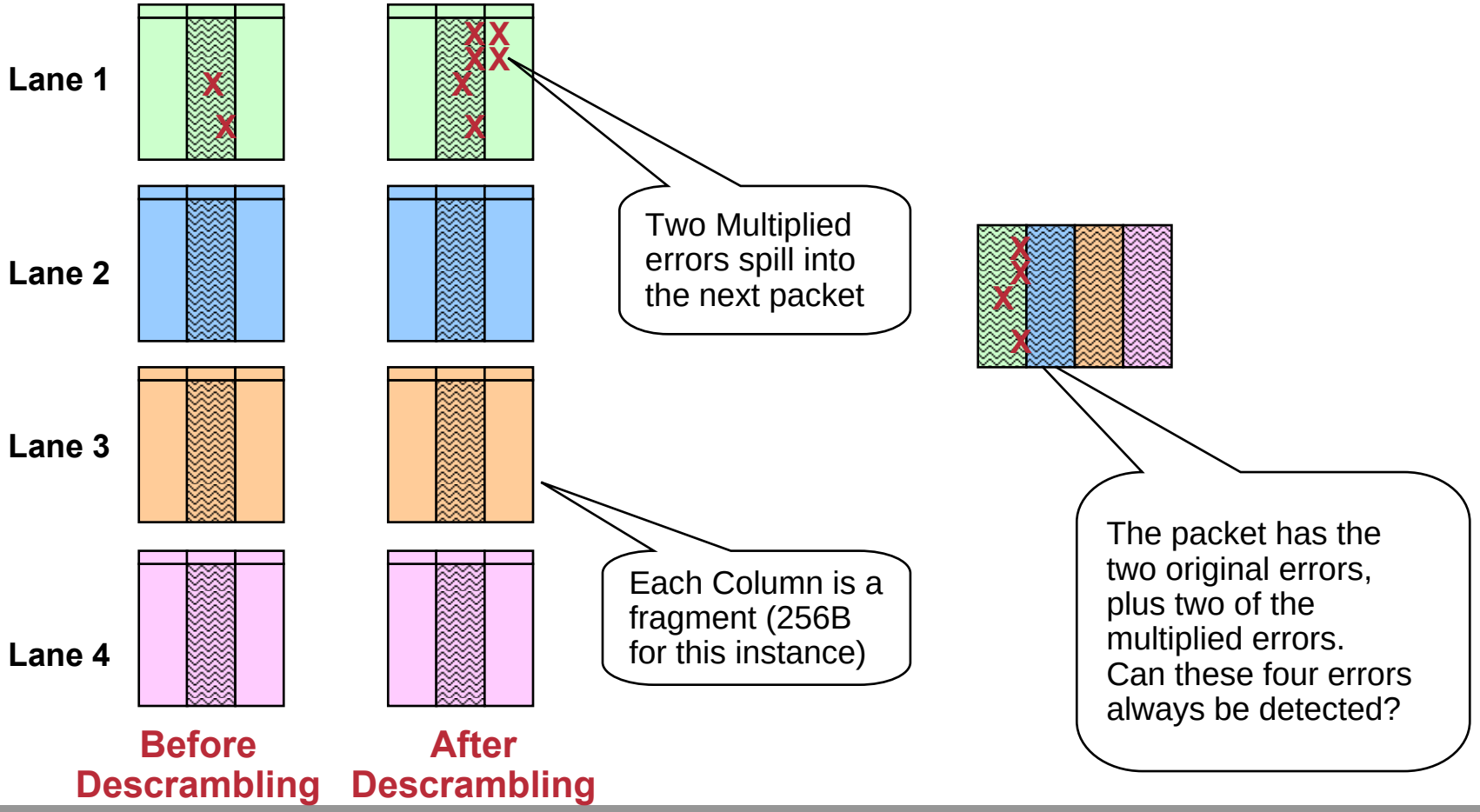
Scrambling at the Electrical Lane Level

- When data is scrambled and descrambled at the electrical lane level (40G or 100G), then the MTTFPA analysis is more complicated
- Example: Two bit random errors (8B data striping example, 4 lanes):



Scrambling at the Electrical Lane Level

- When data is scrambled and descrambled at the electrical lane level (40G or 100G) then the MTTFPA analysis is more complicated
- Example: Two bit random errors (256B data striping example, 4 lanes):



Scrambling at the Electrical Lane Level

- When data is scrambled and descrambled at the electrical lane level (40G or 100G), then the MTTFPA analysis is more complicated
- This is very different than 10GBASE-R, errors often spill out of the fragments, regardless of the fragment size
- Multiplied errors are not necessarily 58 and 39 bit away from the original errors, and therefore we can't automatically say that they are detectable
- All of these cases have to be analyzed
- The above is true regardless of the fragmentation size, 8B, 16B, 256B etc…
- Similar cases occur if you scramble at the optical lane level and virtual lane level

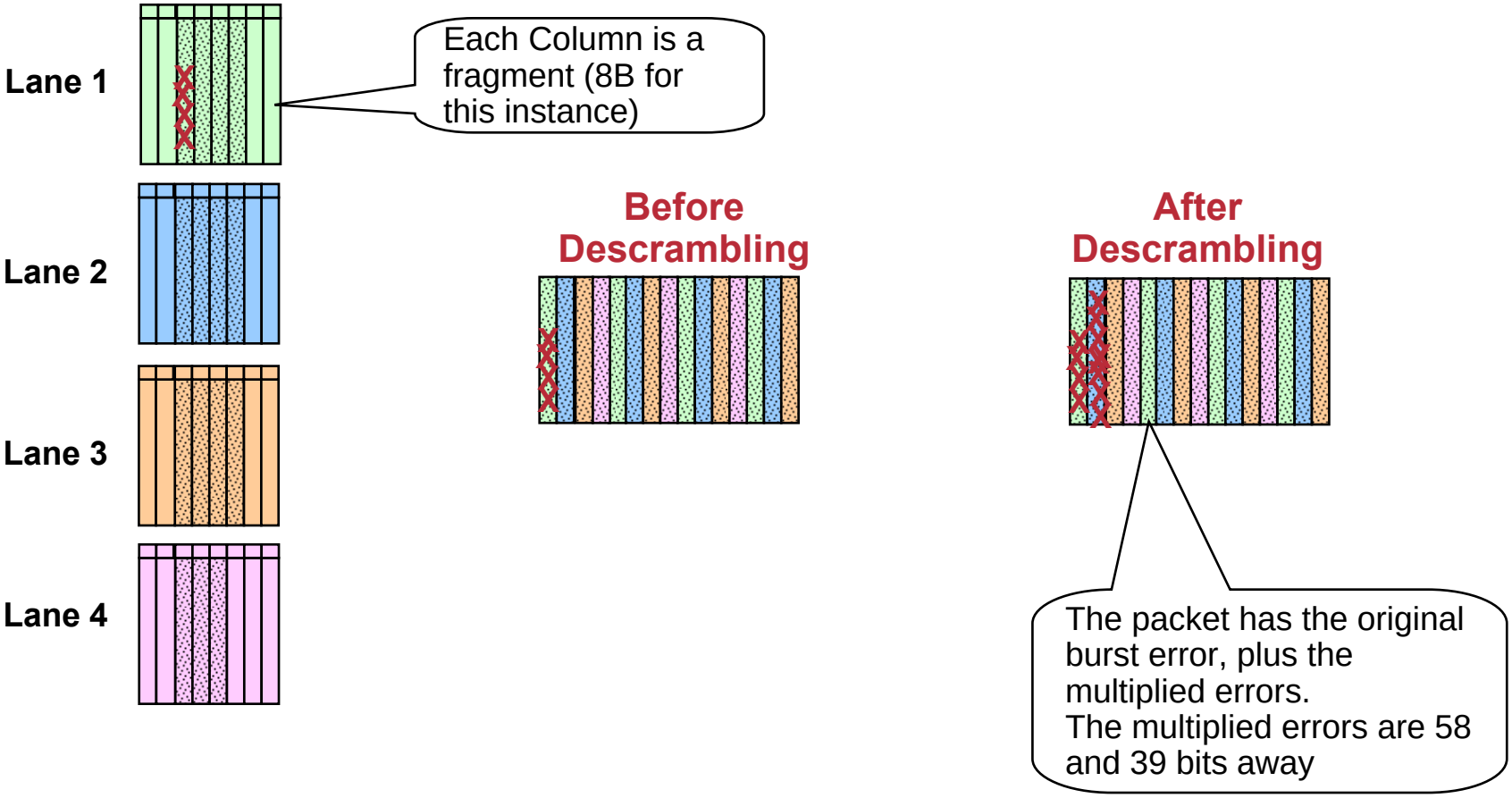
Burst Errors

What about burst errors?

- For 10GBASE-R single burst errors up to 32 bits are always detected by CRC32
- Double burst errors up to 8 bits each are also always detected
- If scrambling is done at the aggregate level, in many cases it still looks like 10GBASE-R
- Other scrambler locations require more complicated analysis

Scrambling at the Aggregate Level

- When data is scrambled and descrambled at the aggregate level (40G or 100G) then the MTTFPA analysis is simpler
- A 4 bit burst error (8B data striping example, 4 lanes):

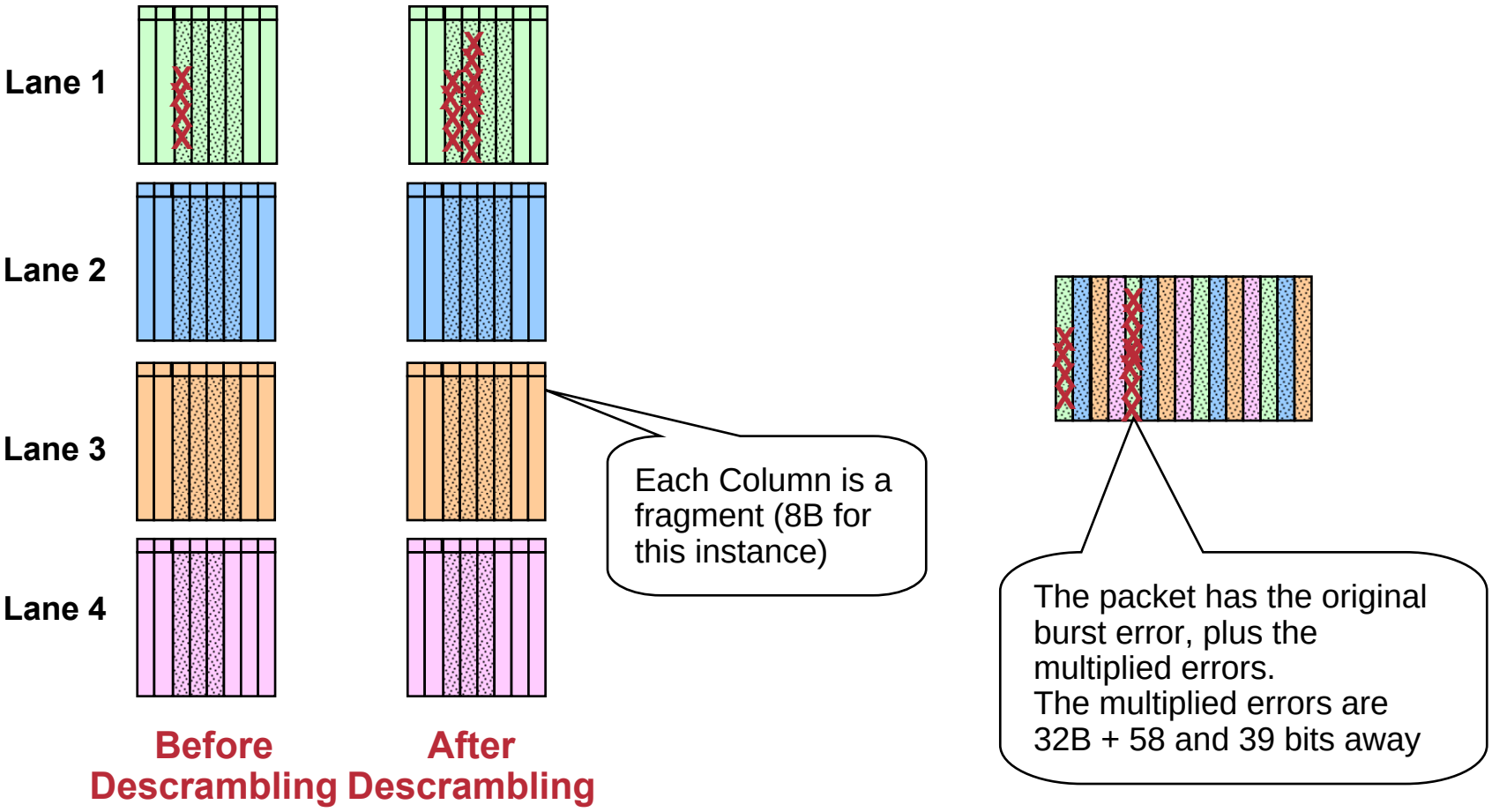


Scrambling at the Aggregate Level – Burst Errors

- When data is scrambled and descrambled at the aggregate level (40G or 100G) then the MTTFPA analysis is similar to 10GBASE-R
- Looks just like the 10GBASE-R case, if the original error is detectable, then the multiplied errors are detectable
- Some boundary cases still have to be analyzed, for instance burst errors that cross the fragmentation boundaries
- The above is true regardless of the fragmentation size, 8B, 16B etc…

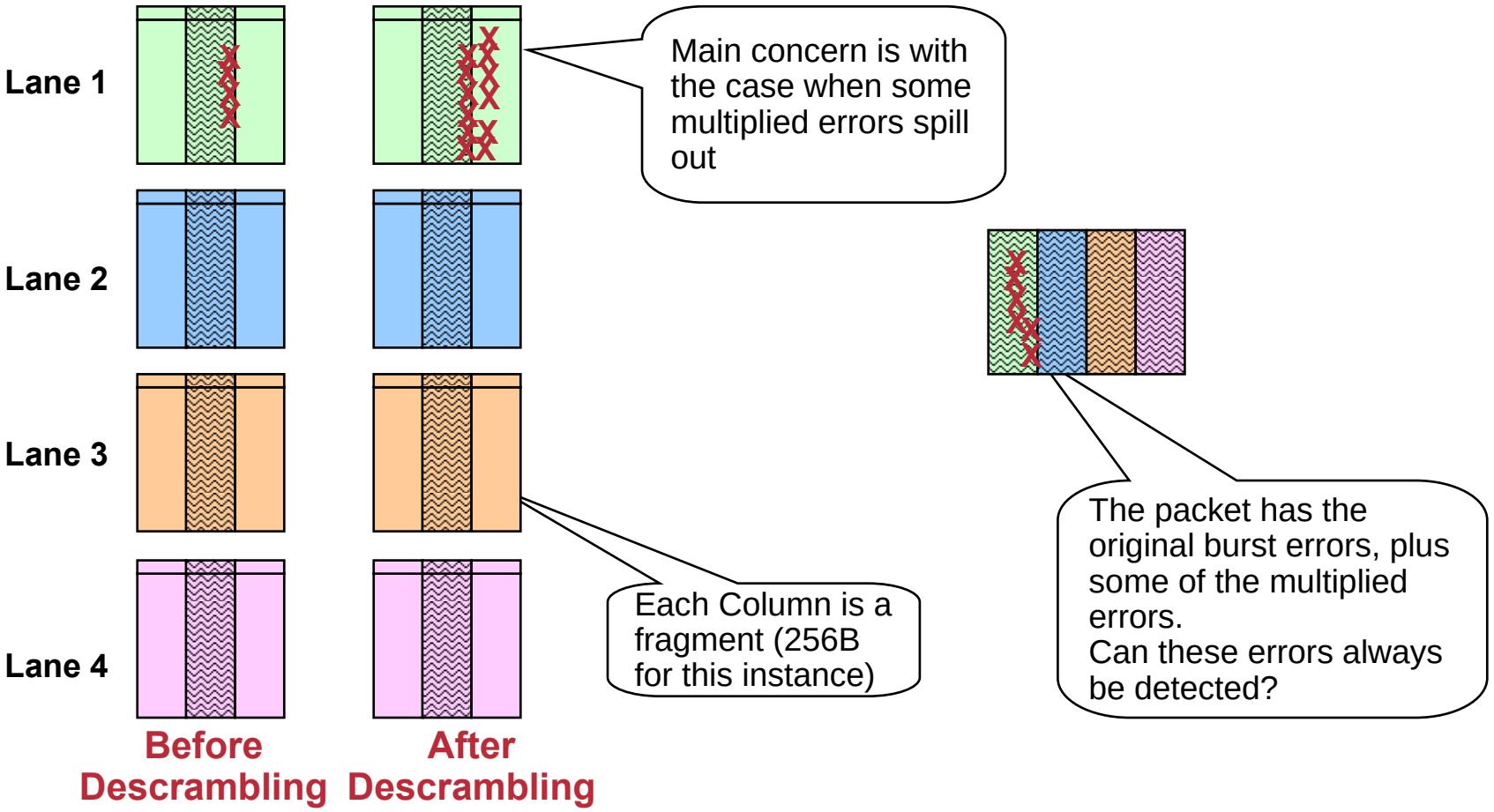
Scrambling at the Electrical Lane Level

- When data is scrambled and descrambled at the electrical lane level (40G or 100G) then the MTTFPA analysis is more complicated
- A 4 bit burst error (8B data striping example, 4 lanes):



Scrambling at the Electrical Lane Level

- When data is scrambled and descrambled at the electrical lane level (40G or 100G) then the MTTFPA analysis is more complicated
- A 4 bit burst error (256B data striping example, 4 lanes):



MTTFPA Summary and Next Steps

MTTFPA

- Scrambling and descrambling data at the 100G or 40G aggregate level greatly simplifies the MTTFPA analysis
  - Looks just like 10GBASE-R in most cases
  - Some boundary conditions still to be analyzed
- Scrambling on a per lane basis complicates the MTTFPA analysis
  - Many more spill in and spill out cases to analyze

Burst Errors impact on MTTFPA

- Need to analyze the likelihood of burst errors on the electrical interface
- Continue to analyze the likelihood of burst errors on the optical interface
- Analyze the boundary conditions of errors crossing sync fields
- Once we have probabilities for different burst errors sizes we can calculate the MTTFPA including this information
- A guarantee of 32 bit burst error detection *may not be possible.* This is not a requirement but has always been supported.