MTTFPA Consideration in Cut Through Mode

Tongtong Wang, Norman Finn, Xinyuan Wang
Huawei Technologies

HUAWEI TECHNOLOGIES CO., LTD.

Cut Through Operation

- What happens in bridge for cut-through operation?
  - Ingress port check packet header (maybe for different depth)
  - Set up temporary Cut-through forwarding path
  - Put chunks of packet in cut-through FIFO instead of Store and Forward FIFO;
  - Block other concurrent transmissions on egress port;
- Previous slides talk about limitation and scenarios for Cut through
  - new-tsn-thaler-cut-through-issues-0117-v01.pdf
  - new-woods-cutthroughconsiderations-0518-v01.pdf
  - Cut through is likely on same speed port connections
  - In industry scenario , typically with chains of two port bridge;
  - Need to minimize the harm of BER in header, constraints on two port device and Qci can help minimize this problem;
- This contribution presents our understanding MTTFPA and suggestions to protect MTTFPA in cut-through mode;

MTTFPA（Mean Time To False Packet Acceptance）

- “MTTFPA is also known as Undetected Error Rate, it is a basic requirement of IEEE 802.3 standard”，*law_01_0907.pdf*
- Original MTTFPA calculation method in consensus is based on the analysis in *walker_1_0300.pdf*
- MTTFPA evaluation is valid from GE, 10GE to 100GE/400GE.
  - For those high quality physical links, with lower speed or shorter reach, 4-bit hamming distance in 802.3 8/10B or 64/66B encoding and CRC32 is enough to ensure MTTFPA close to Age Of Universe (1e10 years)
  - For 100GE/400GE, some longer reach physical links needs FEC(forward error correction) module to boost system performance against error prone physical medias.

MTTFPA without FEC in MAC/PHY

- Dominant errors are spontaneous and randomly distributed.

False Packet Acceptance Rate

- P = coded packet size = 58+1526*8*66/64
- p_e = bit error rate, N = number of errors, t_b = bit time (1/10.3125G).
- probability of N errors in packet of size P:

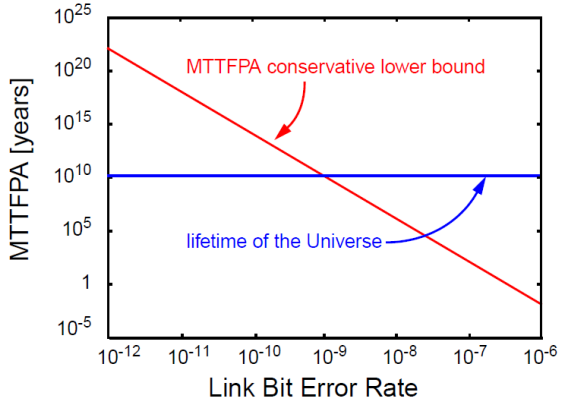
$$p(N, P, p_e) = (1 - p_e)^{P-N}(p_e)^N\binom{P}{N}$$

- expected time for 4 or more errors:

$$MTTFPA > \frac{t_{bit}P}{1 - p(N, P, 0) - p(N, P, 1) - p(N, P, 2) - p(N, P, 3)}$$

$$MTTFPA > \frac{t_{bit}P}{1 - p(0, P, P_e) - p(1, P, P_e) - p(2, P, P_e) - p(3, P, P_e)}$$

False Packet Acceptance Rate



http://grouper.ieee.org/groups/802/3/ae/public/mar00/walker_1_0300.pdf

MTTFPA with FEC in MAC/PHY

- Walker's equation is not valid within this scenario,
- MTTFPA performance is mainly guaranteed by FEC capability.

Undetected FEC Block Error Probability

- If there are t=floor((n-k)/2) or less symbol errors in a FEC block, RS decoder corrects all errors in a FEC block.
- If more than t symbol errors in a FEC block occur, there are two possible cases:

  1) RS decoder fails to decode FEC block and reports to MAC layer that retransmission is required.

  2) RS decoder makes an undetected error because it finds a legal code word other than the transmitted code word. Most undetected FEC block errors are usually detected by inverse transcoder or CRC checker in MAC layer and therefore result in retransmission of an erroneous MAC frame.
- Dominant term in the probability of undetected FEC block errors P_{un}

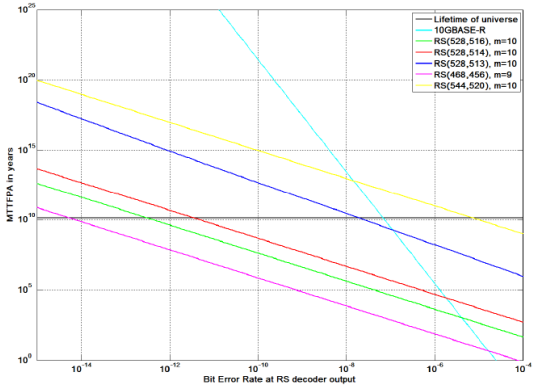
$$P_{un} \approx \left(2^m - 1\right)^{-(d-t-1)} \binom{n-d+t}{t} \Pr\{d-t \text{ symbol errors}\}$$

- For the RS(528,513) m=10 code: $P_{un} \approx 1.61 \times 10^{-9}$ Pr{ 9 symbol errors }
- For the RS(528,514) m=10 code: $P_{un} \approx 1.67 \times 10^{-6}$ Pr{ 8 symbol errors }

IBM MTTFPA Estimation for 100 Gb/s Backplane and Copper Cable 4

MTTFPA Estimate including 2^{32} CRC Factor

- MAC Frame size = 1280 bytes = payload of 2 FEC Blocks
- MTTFPA estimate multiplied by CRC detection factor $2^{32} \approx 4*10^9$
- Working assumption of 2^{32} CRC factor valid if the probability of undetected error patterns at the input of CRC checker that map into the zero syndrome is $1/2^{32}$ of the probability of all undetected error patterns at the input of CRC checker

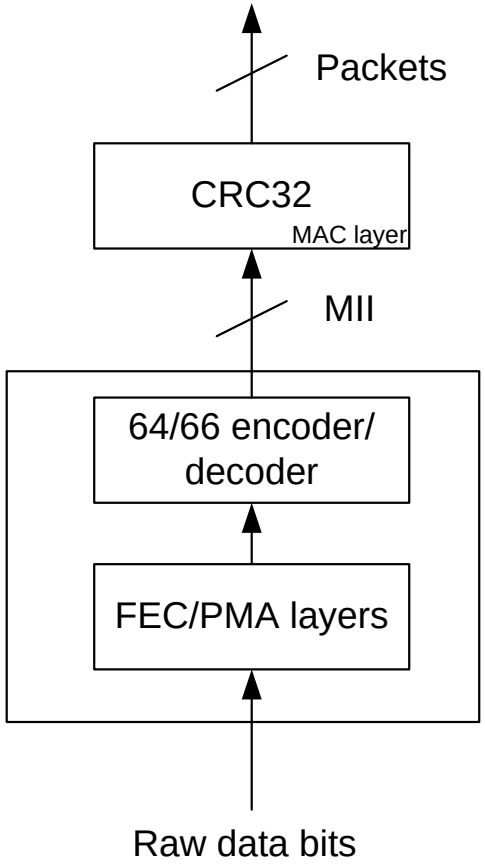


IBM MTTFPA Estimation for 100 Gb/s Backplane and Copper Cable 9

http://www.ieee802.org/3/bj/public/jan12/cideciyan_01_0112.pdf

MII Encoder/Decoder Helps MTTFPA

- From the error detection perspective, CRC32 and 64/66B decoder work together to ensure MTTFPA close to AOU.
  - CRC32 only detects errors inside a packet, checks received payload and checksum with a certain polynomial.
  - 64/66B encoder/decoder help check packet boundary; If packet boundary is wrong, CRC calculation is not accurate and can not ensure reliable transmission.
- Current MTTFPA evaluation is based on CRC32 capability and assumes errors on packet boundary is protected by 64/66B encoder.
- A standard bridge discards bad frames, and relays the CRC-32, to maintain multi-hop protection very close to one-hop protection.



Cut-through Degrades Error Detection

- Today, with cut-through (a frame with a bad CRC is not discarded), 32 hops gives 32 chances to get a 4-bit error. **32 hops = 32 times higher chance of accepting a failed packet**.
- Errors on each hop accumulate. Effectively, 32 hops multiplies the **bit error rate** by 32. If you look at the chart from the Walker slides (slide 4), you see that this **increases the chances of accepting a failed packet many orders of magnitude, not just by a factor of 32**.
- Also, cut-through forwarding does not pass **64/66 encoding errors** to the output port. This further increases the error rate, because Tx port re-generate packet boundary for the error packet and compromise the error detect capability of the end node. This increases failed packet acceptance much faster than linearly with the number of hops.

Possible Corrections

There are a number of possible steps we can take to restore the current MTTFPA. 802.1 needs to explore these with 802.3.

- Do not lose the encoding errors. That is, provide some means for saying, “This frame is bad,” to the output MAC before the last bit has been transmitted.
- A frame so labeled must reliably be passed on as bad at each hop; the label cannot easily be transformed into a good frame.
- Use 802.1CB to make up for increased bit error rate.
- Employ FEC end-to-end above the MAC.
- Count labeled-bad frames separately from CRC-32 errors (to prevent mistaking bad-frame labels from eye diagram errors).

Thank you