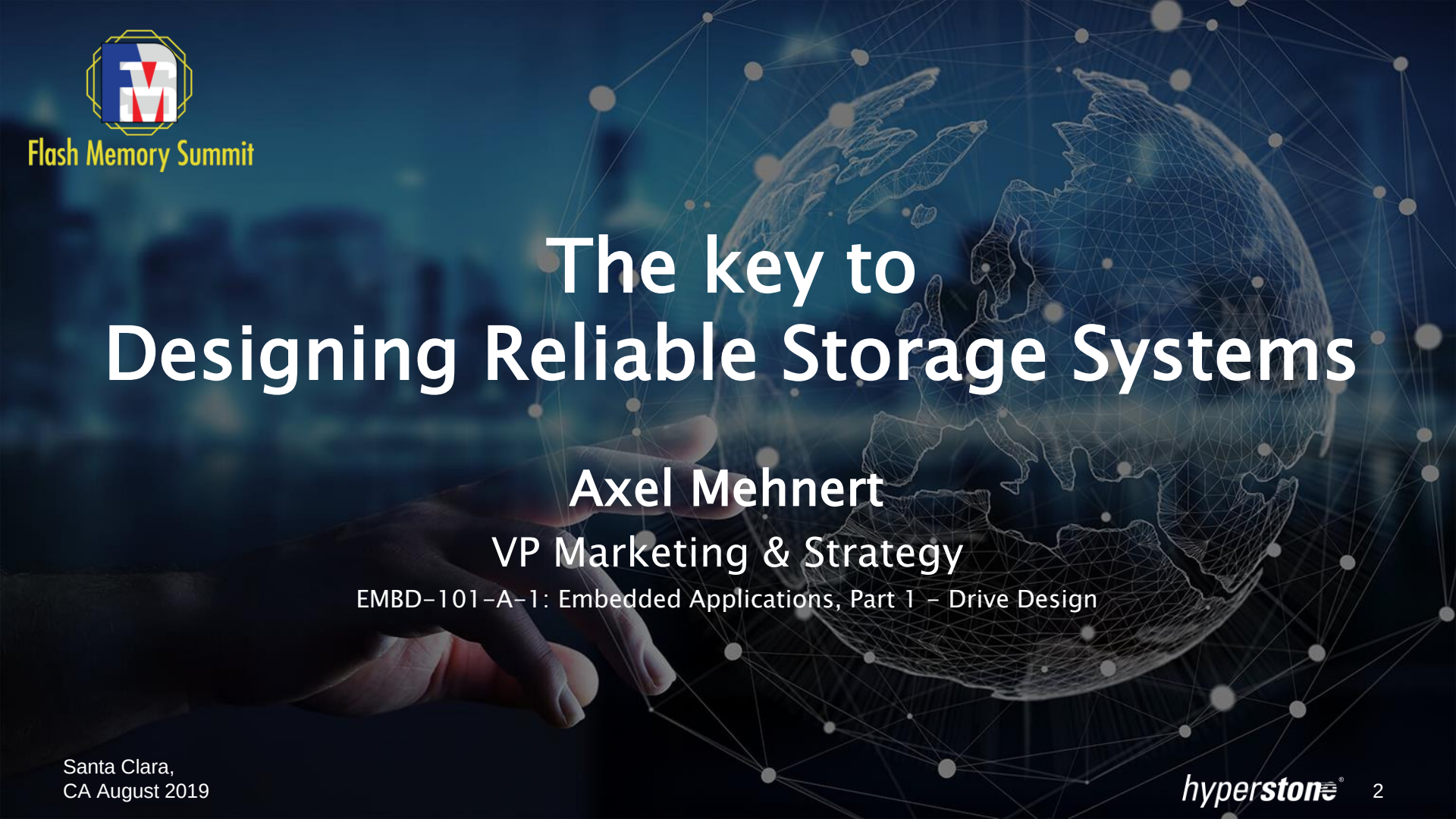
"The bitterness of poor quality remains long after the sweetness of low price is forgotten."

– Benjamin Franklin –



The key to Designing Reliable Storage Systems

Axel Mehnert

VP Marketing & Strategy

EMBD-101-A-1: Embedded Applications, Part 1 – Drive Design

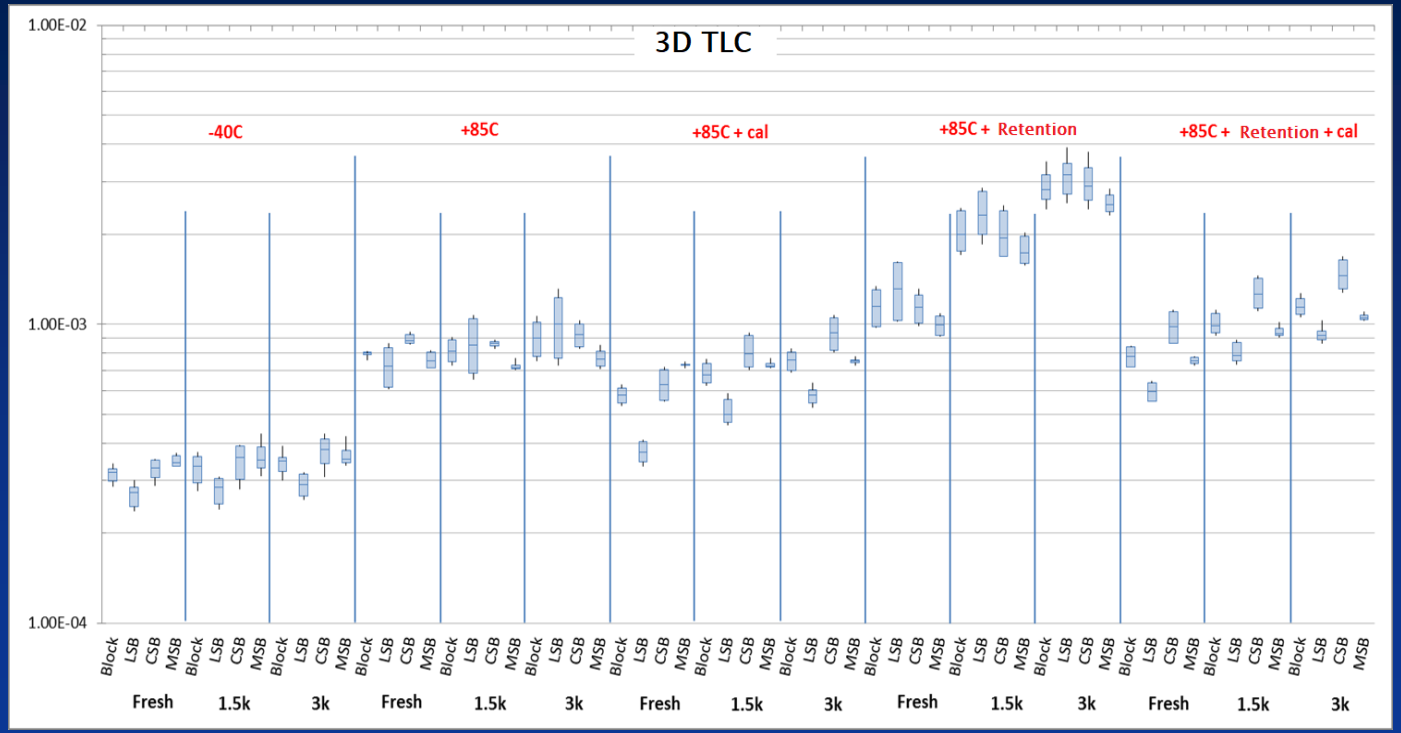
Santa Clara,
CA August 2019

System Level Reliability

- Some quality needs to be by-design

- Our design and reliability target:
  Highest correction performance with guaranteed error floor
  - UBER at no less than $1x10^{-16}$ worst case

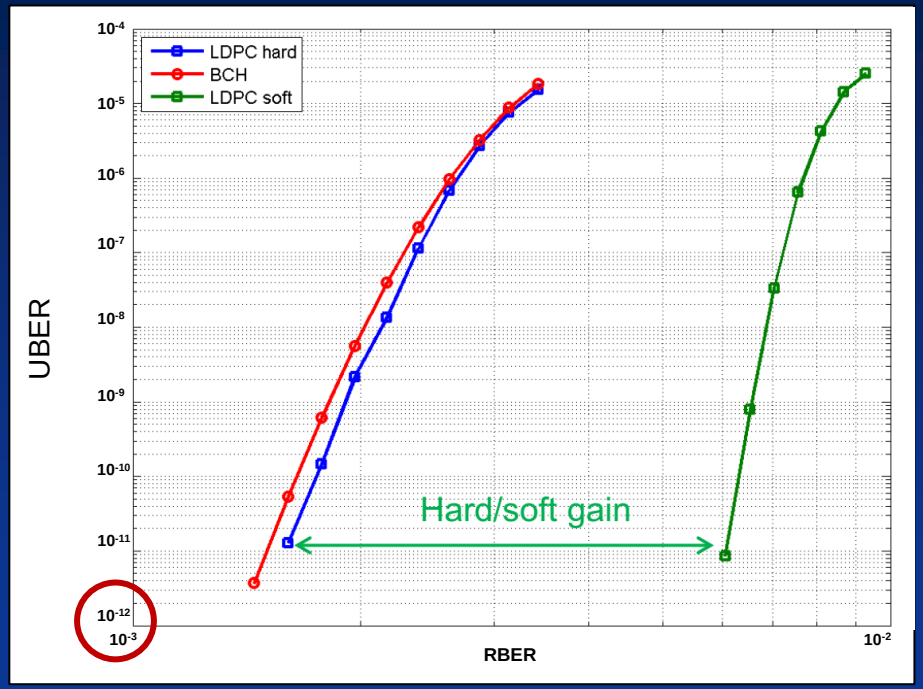
Lifetime RBER



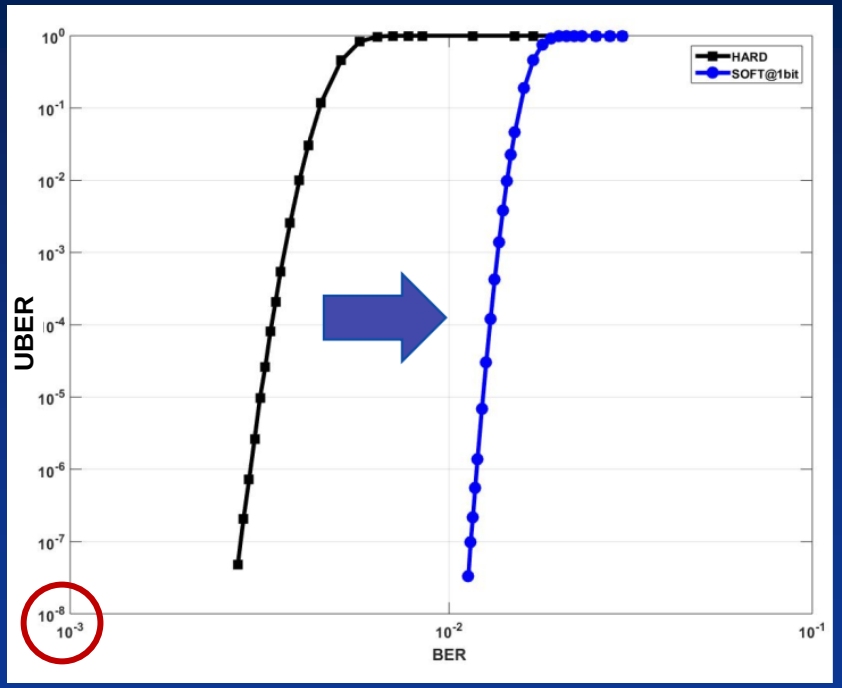
Cross-temp program at -40° C read at different temperatures

After calibration, RBER between $2x10^{-4}$ and $2x10^{-3}$ (EOL & worst case)

Typical ECC performance plots

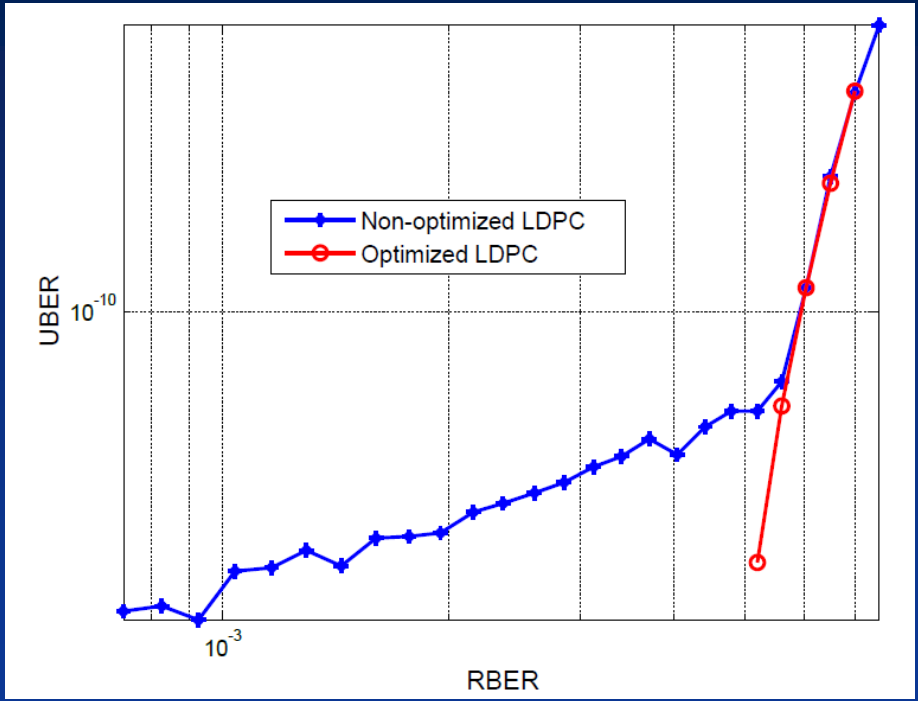


"LDPC Code Concepts and Performance on High-Density Flash Memory", Erich F. Haratsch, Flash Memory Summit 2014

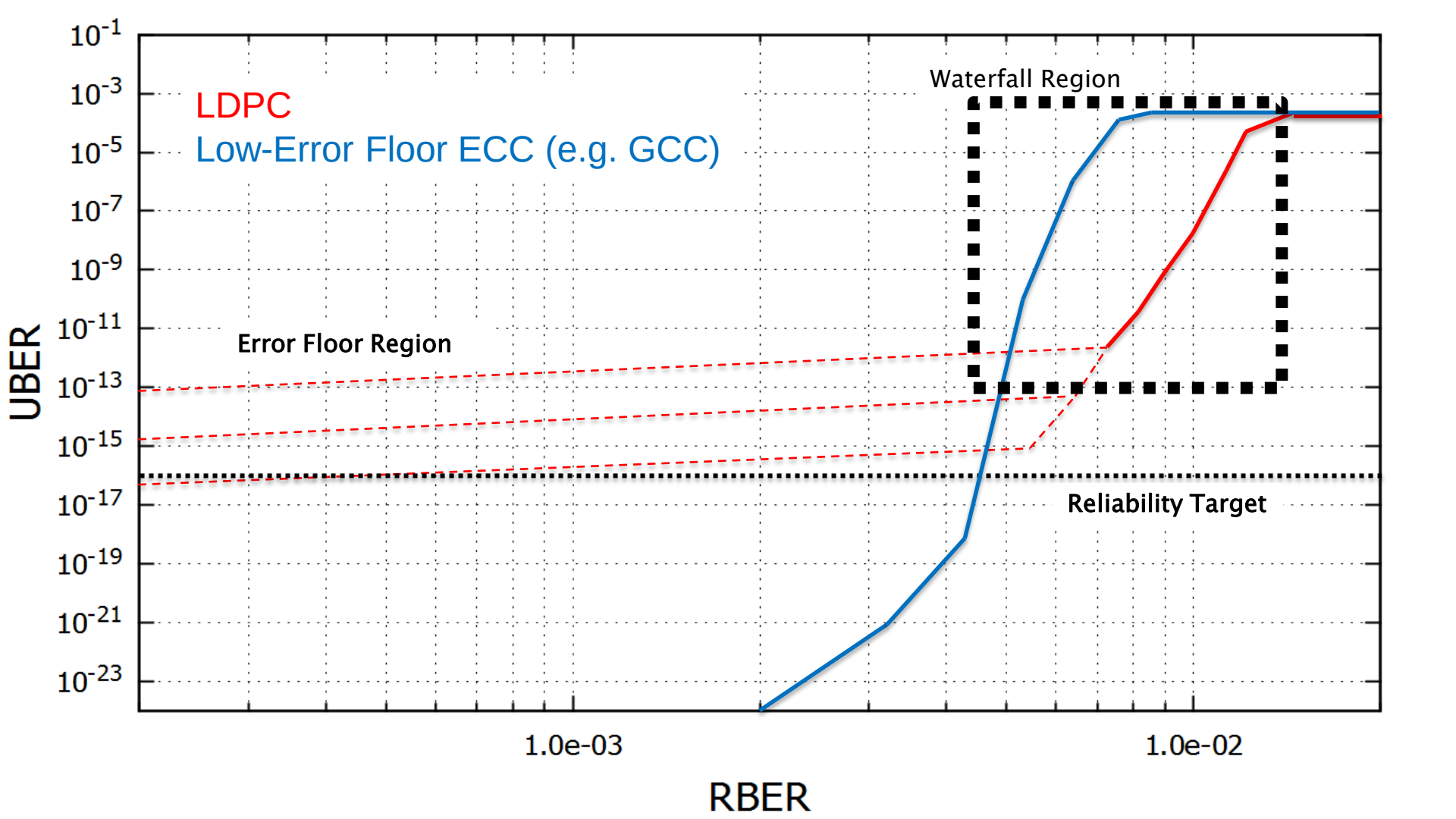


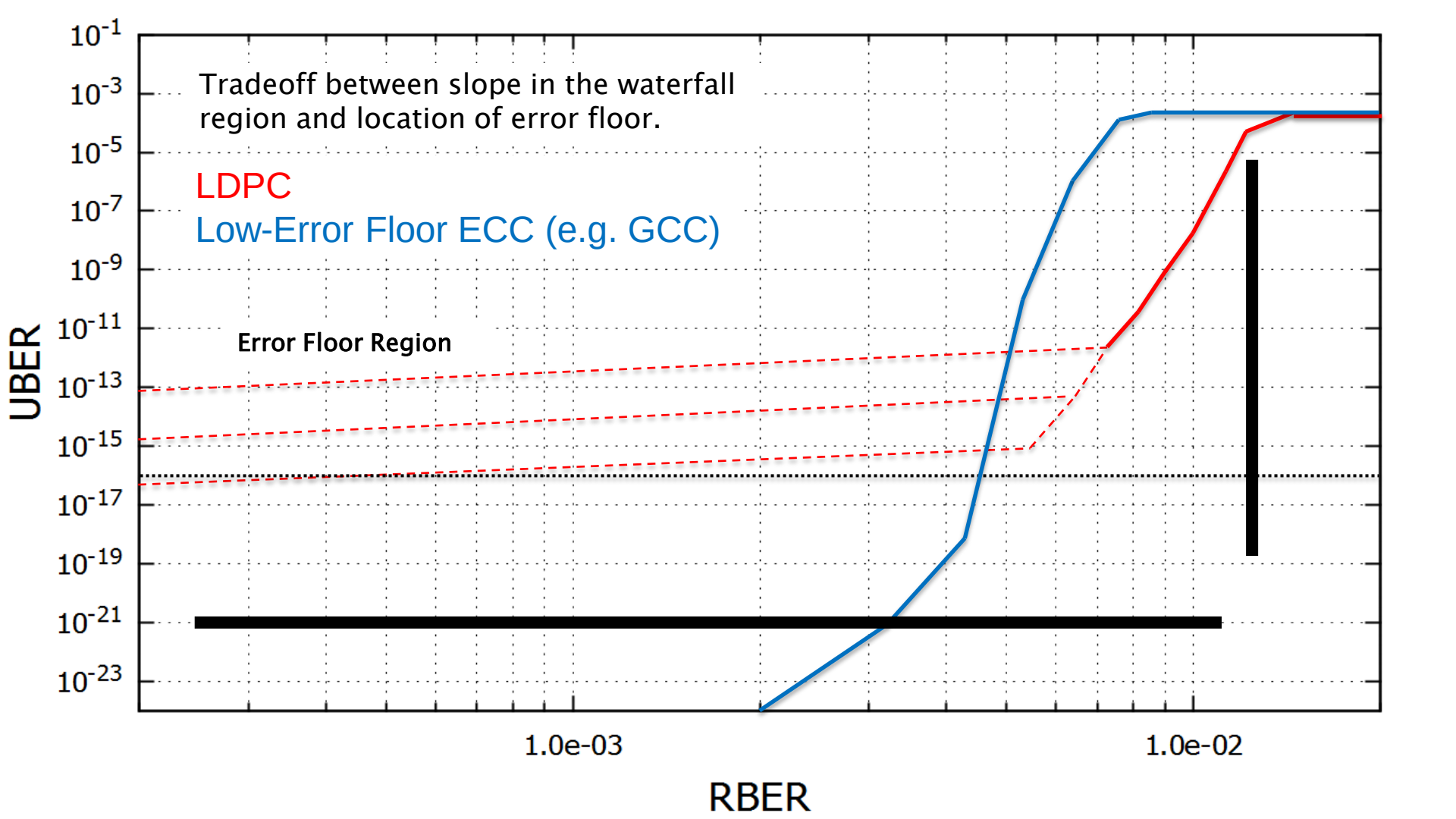
"Fully Integrated LLR Calculation Flow", L. Zuolo, A. Marelli, and R. Micheloni, Flash Memory Summit 2018

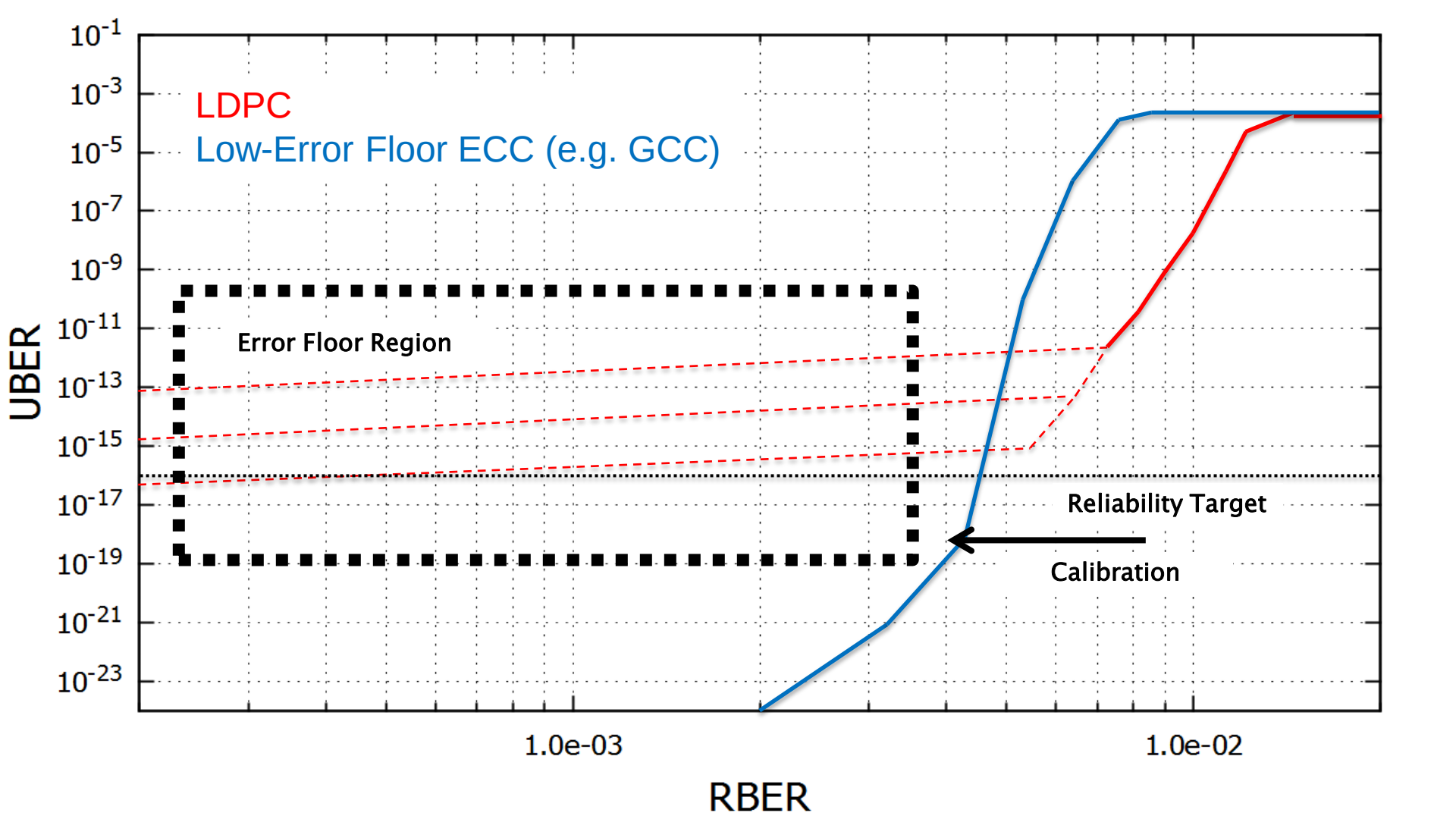
Where exactly is the error floor of the ECC you are using?

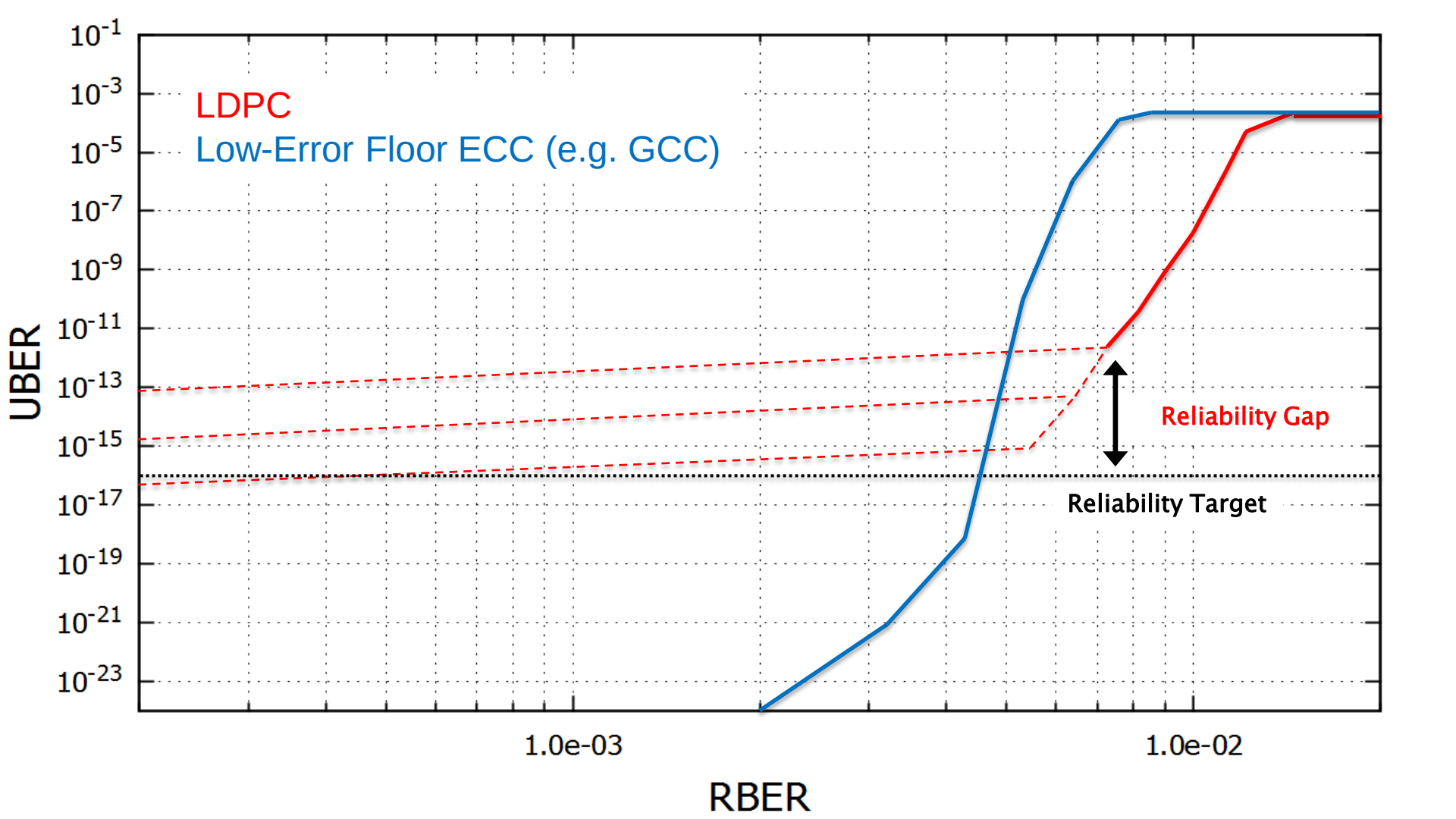


"LDPC Code Concepts and Performance on High-Density Flash Memory", Erich F. Haratsch, Flash Memory Summit 2014


LDPC
Low-Error Floor ECC (e.g. GCC)
Waterfall Region
Error Floor Region
Reliability Target
UBER
RBER
10-1
10-3
10-5
10-7
10-9
10-11
10-13
10-15
10-17
10-19
10-21
10-23
1.0e-03
1.0e-02


Tradeoff between slope in the waterfall region and location of error floor.
LDPC
Low-Error Floor ECC (e.g. GCC)
Error Floor Region
UBER
RBER
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
10^{-13}
10^{-15}
10^{-17}
10^{-19}
10^{-21}
10^{-23}
1.0e-03
1.0e-02


LDPC
Low-Error Floor ECC (e.g. GCC)
Error Floor Region
Reliability Target
Calibration
UBER
RBER
10-1
10-3
10-5
10-7
10-9
10-11
10-13
10-15
10-17
10-19
10-21
10-23
1.0e-03
1.0e-02


LDPC
Low-Error Floor ECC (e.g. GCC)
Reliability Gap
Reliability Target
UBER
RBER
10-1
10-3
10-5
10-7
10-9
10-11
10-13
10-15
10-17
10-19
10-21
10-23
1.0e-03
1.0e-02

RAID vs. TempRAID

RAID (permanent)

Block A	Block B	Block C	Block D
User data	User data	User data	Parity

TempRAID

TempRAID	pSLC	pSLC	pSLC	TLC
Temporary Parity	1 page	1 page	1 page	3 pages
	1 page	1 page	1 page	3 pages
	1 page	1 page	1 page	3 pages

Keep pSLC blocks with RAID as backup until TLC block is successfully written

- WAF increase
- Performance impact
- RAID is not always the solution

RAID vs. TempRAID

Protection Mechanism	Capacity Impact	Recovery Time	Protection Strength
RAID (3:1)	High	Medium	High
RAID (127:1)	Medium	Large	Medium
Read-Verify (temporary SLC RAID)	None	Small	High (if low error-floor ECC is used)

Summary

- Pushing ECC performance to higher RBER may introduce a higher error floor
- RAID as countermeasure to higher UBER is costly (WAF, capacity, performance)
- Low-Error-Floor ECC (e.g. GCC)
  - Allows cheaper and reliable countermeasure (Read-Verify)
  - Allows different/less-overhead RAID structures

Final Thoughts…

Your key take away for the summit

- Watch for the error floor
- Is UBER plotted down to 10^{-16} and beyond?
- Don't trust dotted lines – those may be based on assumptions


Flash Memory Summit
“Small things make perfection, but perfection is no small thing”
– Henry Royce –
Santa Clara,
CA August 2019
hyperstone
15

The secret to Designing Reliable Storage Systems

Additional Information

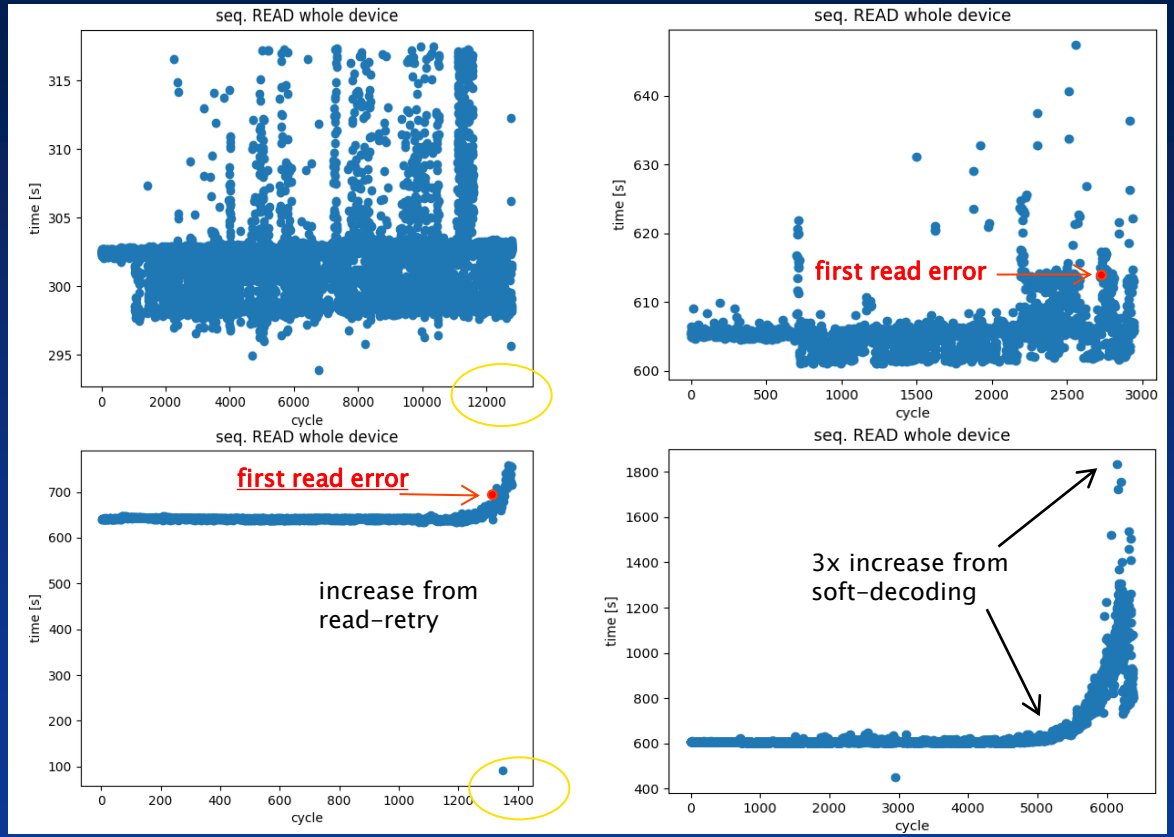
Santa Clara,
CA August 2019

hyperstone®

System Level Reliability Tests

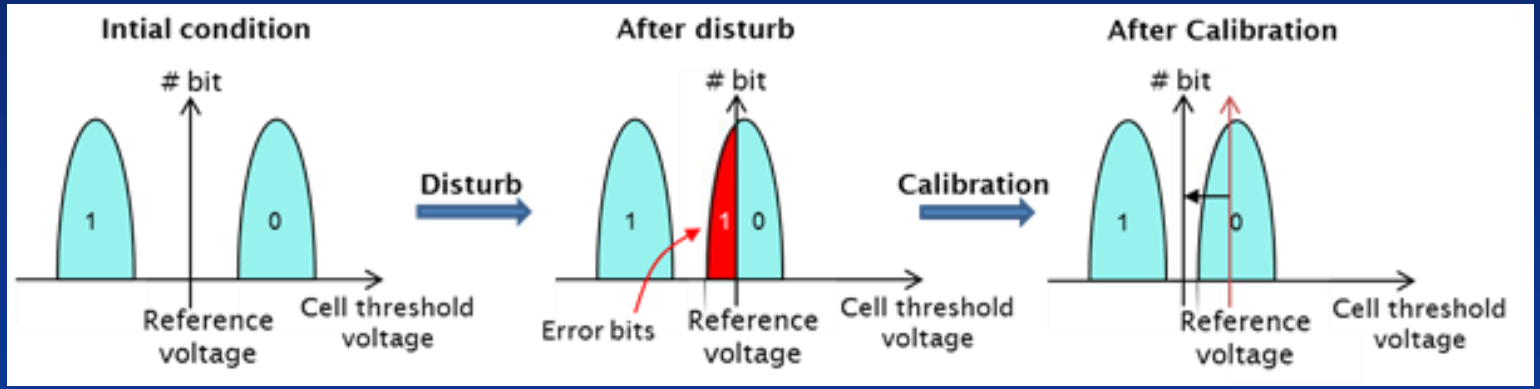
Flash Memory Summit

- Subsequent full drive writes (cycles) / drive reads (drive reads plotted in [s])
- Read errors are tracked, first one is highlighted
- Test stops at first failed write command
- 10x difference in lifetime of different SSDs!

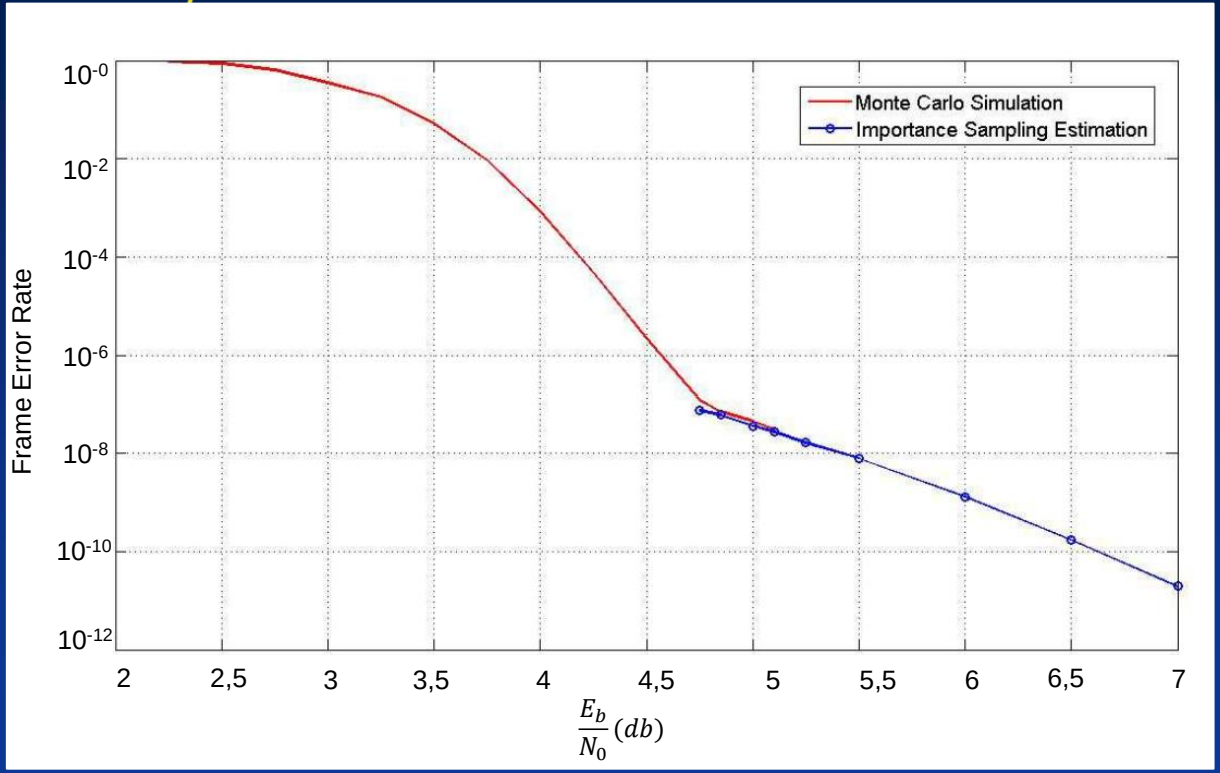


Calibration

Calibration can be carried out by firmware in configurable intervals during operation



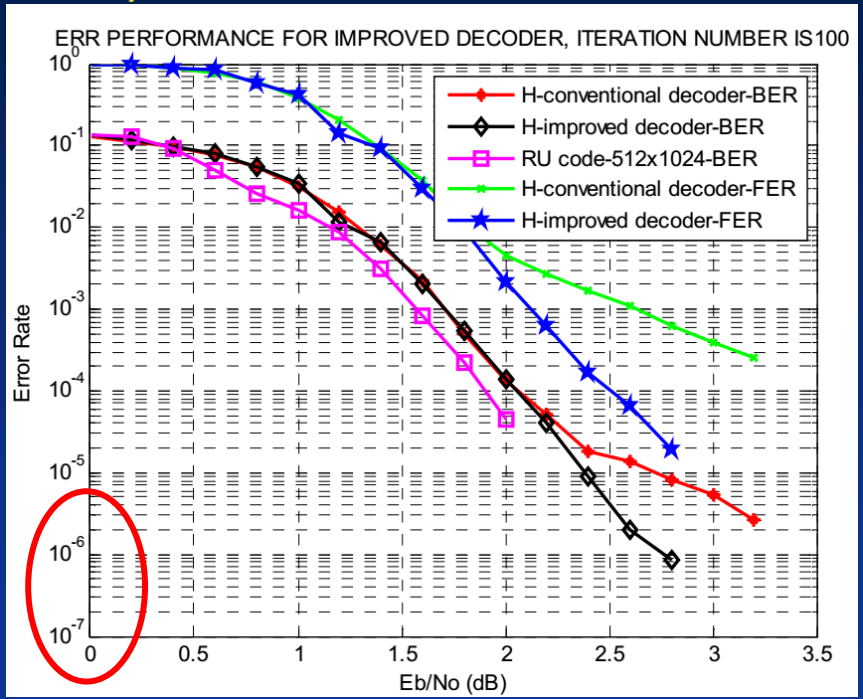
Error Correction



- LDPC codes are not simulated down to JEDEC specified rates (= 10^{-16} for enterprise), but only down to ~10^{-10}
- For the further trend methods like „importance sampling“ are used where a small subset of the codewords is used for estimation
- This educated guess is an orientation but nothing that can be relied on!

Source: “An Efficient Algorithm for Finding Dominant Trapping Sets of LDPC Codes” Mehdi Karimi, Student Member, IEEE and Amir H. Banihashemi, Senior Member, IEEE

Error Correction



"Analysis of Error-Prone Patterns for LDPC Codes under Belief Propagation Decoding ", Huanlin Li, Yanyan Cao, Jeffrey C. Dill, The 2010 Military Communications Conference

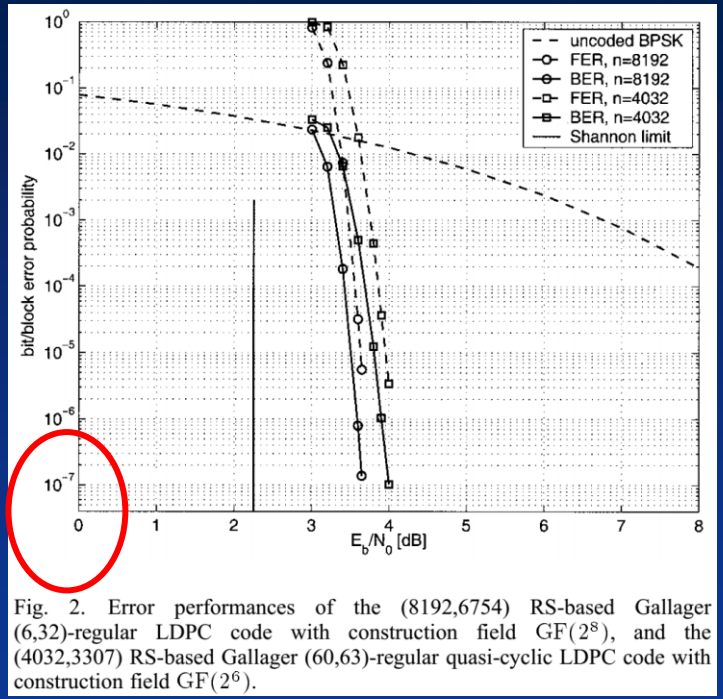


Fig. 2. Error performances of the (8192,6754) RS-based Gallager (6,32)-regular LDPC code with construction field $GF(2^8)$, and the (4032,3307) RS-based Gallager (60,63)-regular quasi-cyclic LDPC code with construction field $GF(2^6)$.

„A Class of Low-Density Parity-Check Codes Constructed Based on Reed-Solomon Codes With Two Information Symbols", Ivana Djurdjevic, Jun Xu, Khaled Abdel-Ghaffar, *Member, IEEE*, and Shu Lin, *Fellow , IEEE*, IEEE COMMUNICATIONS LETTERS, VOL. 7, NO. 7, JULY 2003

Error Correction

- LDPC
  - High performance
  - Error-floor unknown – estimation only (e.g. importance sampling)
  - Full simulation not feasible
  - Many different “quality” grades depending on implementation and alignment to specific Flash technology and channel model

- Generalized Concatenated Code (GCC)
  - High correction performance
  - Error-floor: analytical determination possible
  - Low-Error Floor
  - Guaranteed correction capability

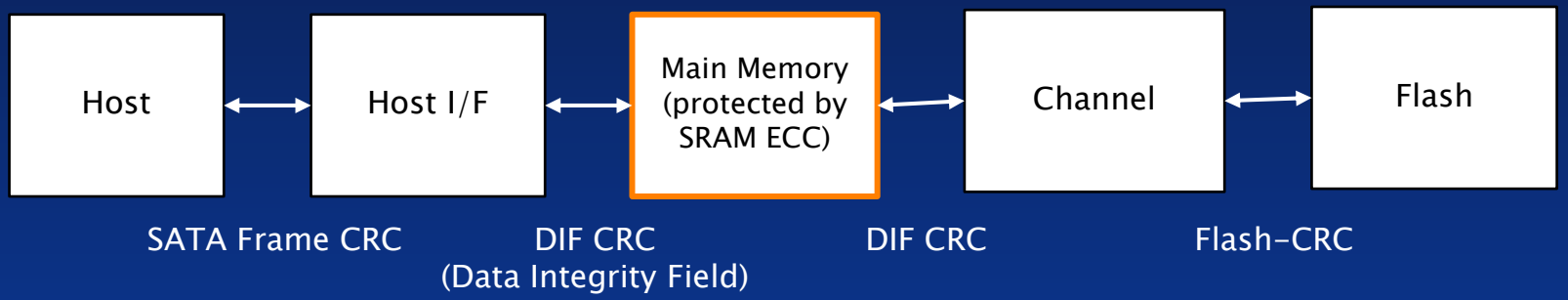
UBER

- Two definitions
  - Scientific: Uncorrectable Bit Error Rate – a statistic measure of the probability of one erroneous bit after processing in the the ECC unit (module level).
  - JEDEC: $UBER = \frac{number\ of\ data\ errors\ (sectors)}{number\ of\ bits\ read}$ during the TBW rating limit of the drive (sector errors of the whole system).
  - Large difference between the two depending on the ECC frame size and system sector size

Reliability Enhancement Techniques

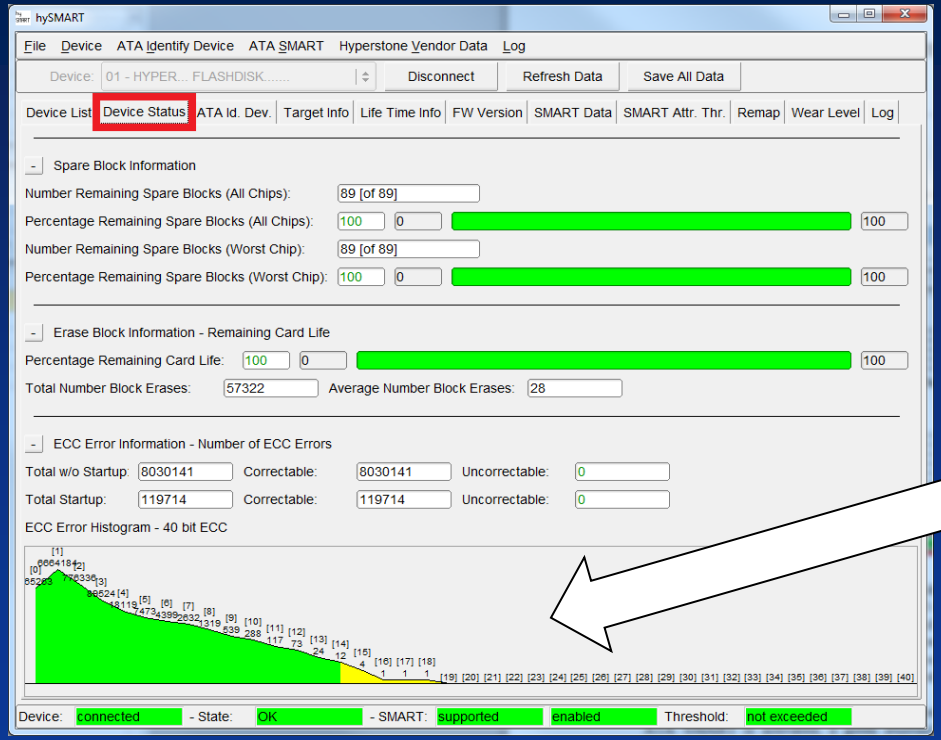
Technique	Impact on Average Read Performance	Impact on Trail Performance (Read)	WAF	Management Overhead
ECC (hard-decision)	Negligible	None	Negligible	None
ECC (soft-decision)	None	High	Negligible	Negligible
Read Retry	None	High	None	Voltage levels
Flash Calibration	Low	High	None	None
RAID (3:1) RAID (127:1)	High (depending on error floor) Low (depending on error floor)	Medium High	High Medium	Medium Low
Temporary RAID (e.g. Read-Verify)	None	None	Medium	Medium
Dynamic Data-Refresh	None	None	Low	Low

E2E Data-Path Protection



- Protection against radiation (single event effects)
- Detection & correction of errors in the main memory
- Comprehensive protection of the complete datapath

RBER – Lifetime Monitoring



With the Hyperstone **hySMART** tool you can monitor the corrected bit errors during lifetime to be confident your system works within operating conditions and to predict lifetime for your specific use-case and mission profile.

hyReliability FlashXE®

FlashXE®

eXtended Endurance – Hardware and Firmware Features

Qualification	Near Miss ECC	Dynamic Data-Refresh
Calibration	Read Retry	Low-error floor ECC
Read Disturb Management	RAID	Read Verify