Flash Reliability in Production: The Expected and the Unexpected

Bianca Schroeder
University of Toronto

Raghav Lagisetty,
Arif Merchant
Google Inc.

Google

Main research interests

- System reliability
- Why and how do systems fail in the wild?

Field data

Data from a large number of large-scale production systems at different organizations:

Field data

- Different hardware failure events
  - Hardware replacements
  - Correctable and uncorrectable errors in DRAM
  - Server outages
  - Hard disk drive failures
  - Sector errors in hard disk drives
  - Data corruption in storage systems
  - Failures in solid state drives
- Job logs
  - Google, OpenCloud (Hadoop cluster at CMU), Yahoo! Hadoop trace

- Observations often different from expectations
- Surprising to operators as well as manufacturers

Field data

- Different hardware failure events
  - Hardware replacements
  - Correctable and uncorrectable errors in DRAM
  - Server outages
  - Hard disk drive failures
  - Sector errors in hard disk drives
  - Data corruption in storage systems
  - Failures in solid state drives

[FAST'16]

Google

- Job logs
  - Google, OpenCloud (Hadoop cluster at CMU), Yahoo! Hadoop trace

- Observations often different from expectations
- Surprising to operators as well as manufacturers

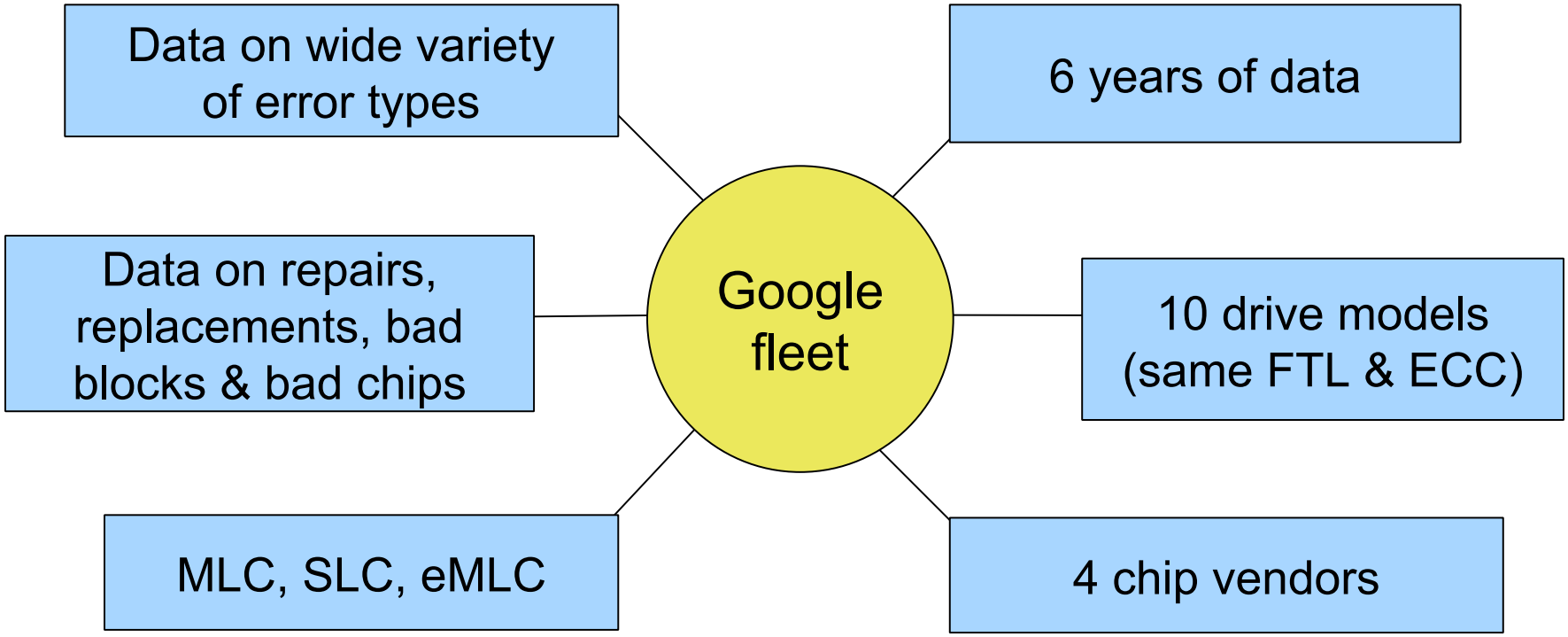
Flash reliability

- Why flash?
  - More and more data is living on flash

    => data reliability depends on flash reliability
  - Worry about flash wear-out

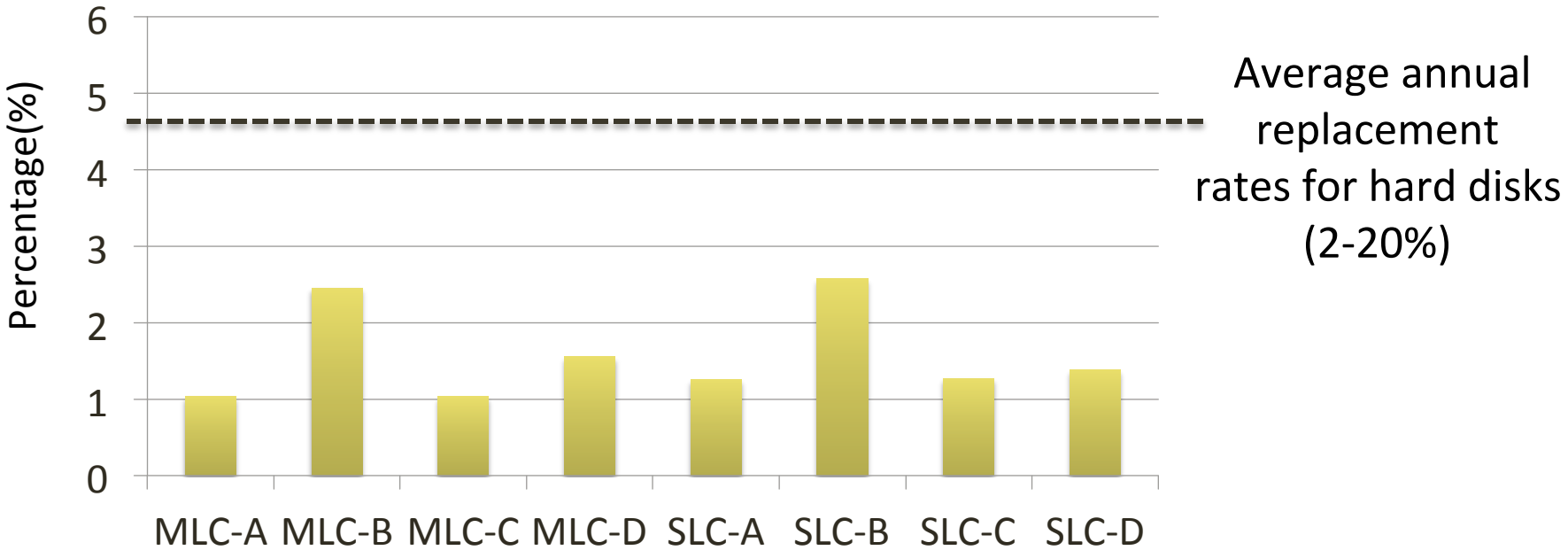
- Little prior work on *production systems*
  - Lab studies using accelerated testing
  - Only one field study (Sigmetrics'15)

The data


Data on wide variety of error types
6 years of data
Data on repairs, replacements, bad blocks & bad chips
Google fleet
10 drive models (same FTL & ECC)
MLC, SLC, eMLC
4 chip vendors

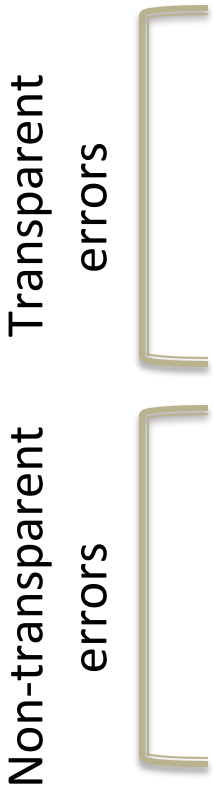
Drive replacements

- Percentage of drives replaced annually due to suspected hardware problems over the first 4 years in the field:

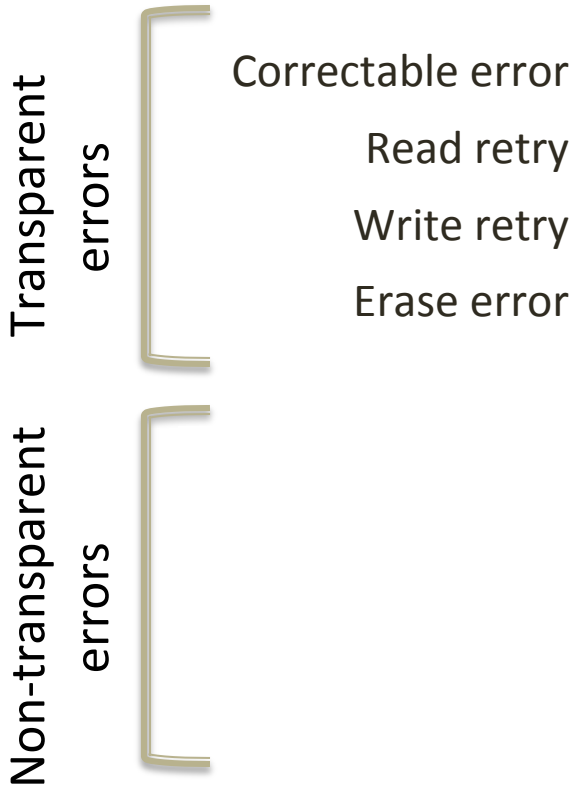


- **~1-2% of drives replaced annually, much lower than hard disks!**
- **0.5-1.5% of drives developed bad chips per year**
  - **Would have been replaced without methods for tolerating chip failure**

Errors experienced during a drive's lifecycle



Errors experienced during a drive's lifecycle



Errors experienced during a drive's lifecycle

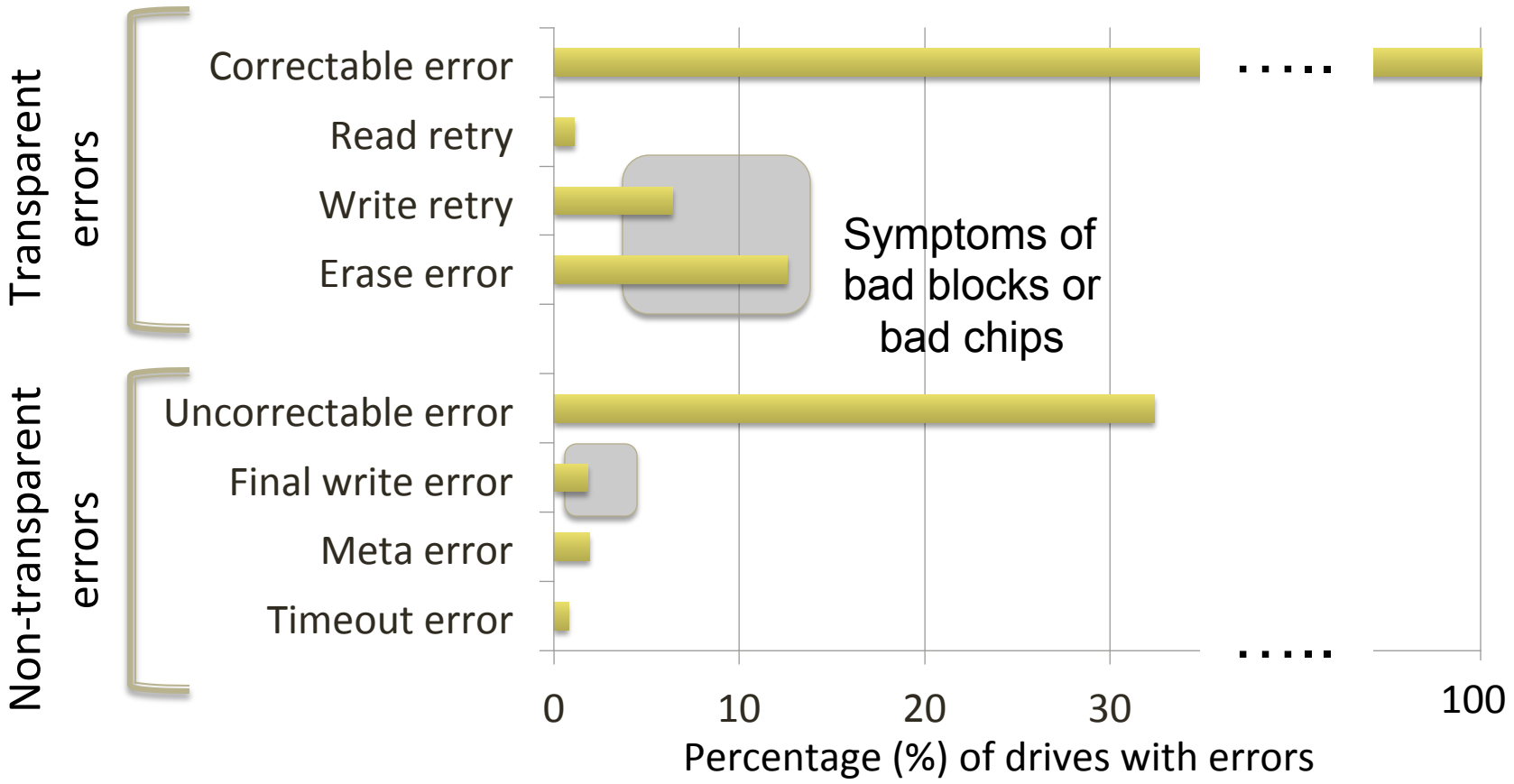
Transparent errors

- Correctable error
- Read retry
- Write retry
- Erase error

Non-transparent errors

- Uncorrectable error
- Final write error
- Meta error
- Timeout error

Errors experienced during a drive's lifecycle



- **Non-transparent errors common:**
  - **26-60% of drives with uncorrectable errors**
  - **2-6 out of 1,000 drive days experience uncorrectable errors**
  - **Much worse than for hard disk drives (3.5% experiencing sector errors)!**

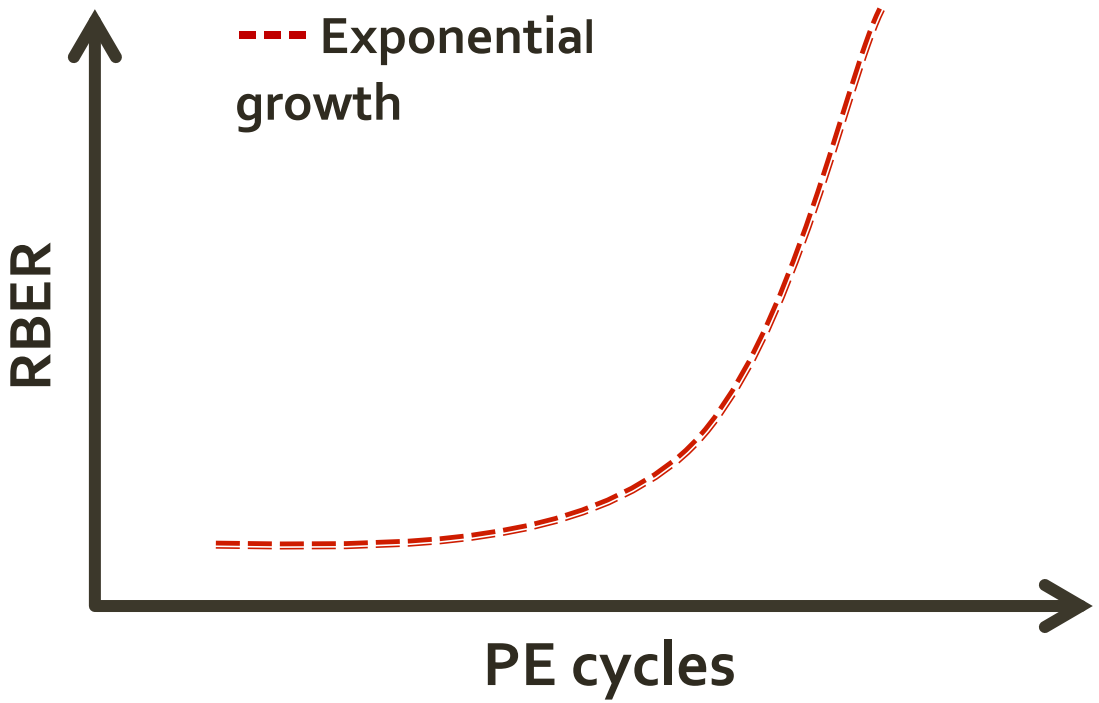
What factors impact flash reliability?

- Wear-out (limited program erase cycles)
- Technology (MLC, SLC)
- Lithography
- Age
- Workload

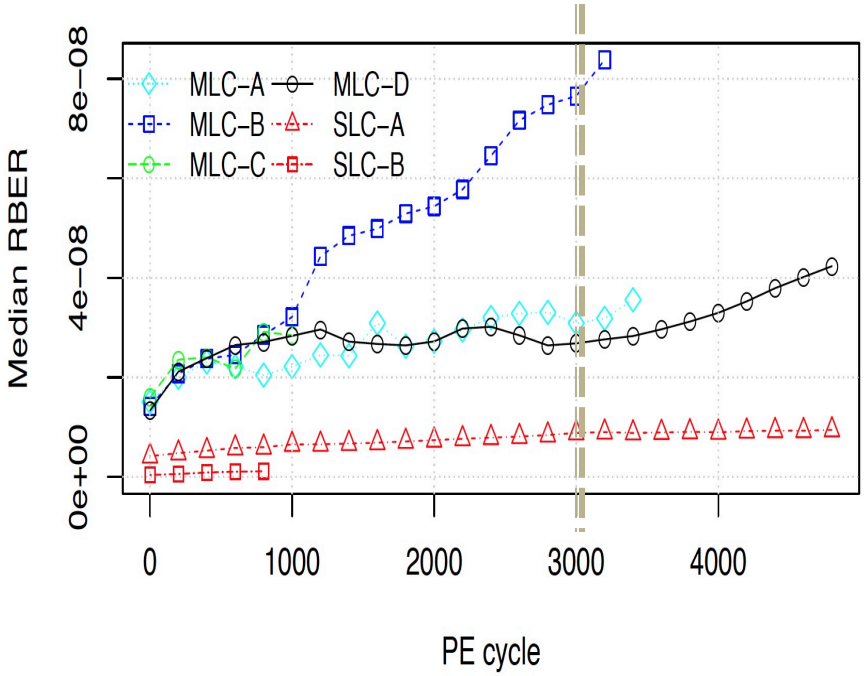
- What reliability metric to use?
  - Raw bit error rate (**RBER**)
  - Probability of **uncorrectable errors**
    - Why not UBER? We shall see …

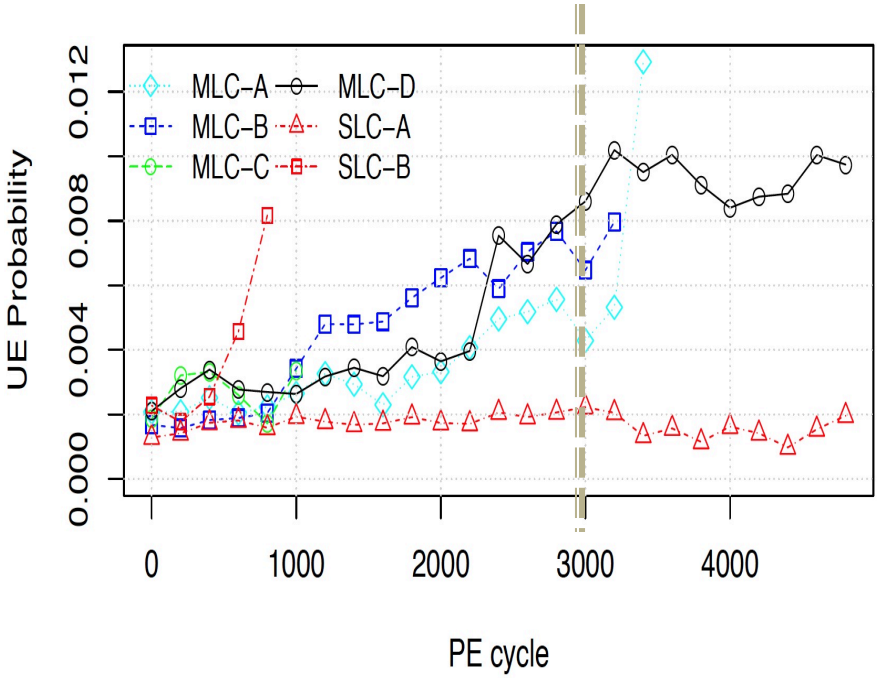
Effect of wear-out (program erase cycles)

Common expectation:
Exponential increase of RBER with PE cycles



Effect of wear-out (program erase cycles)



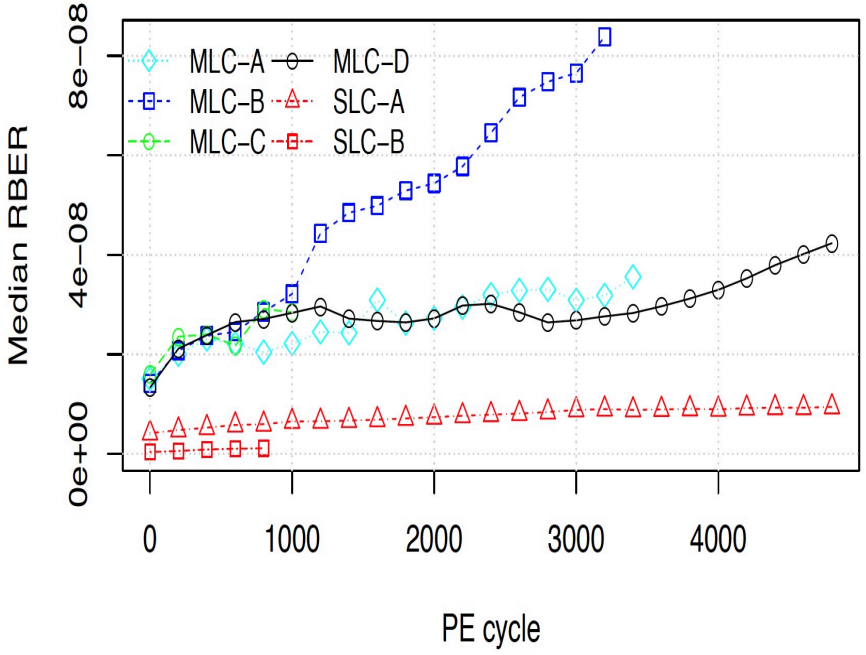


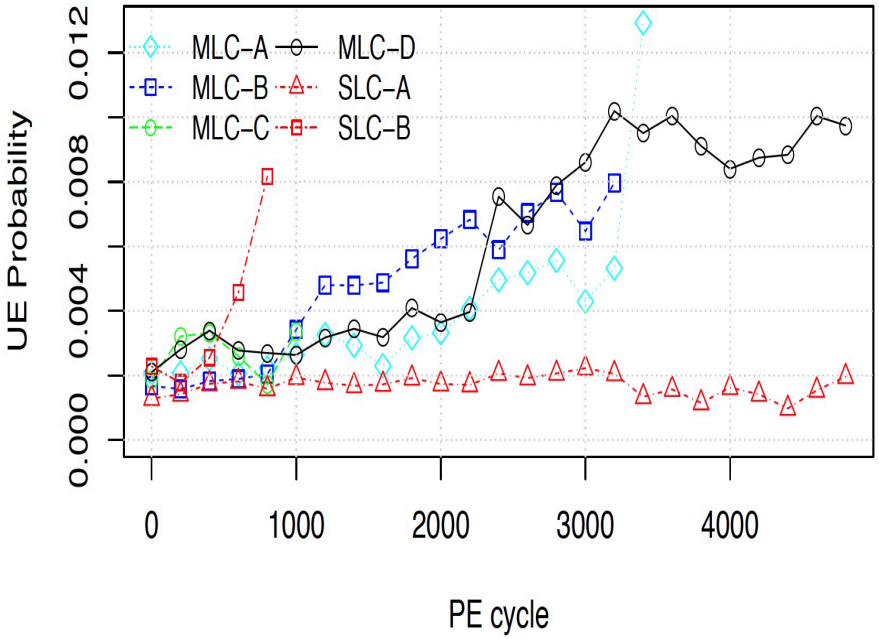
- Big differences across models (despite same ECC)
- Linear rather than exponential increase
- No sudden increase after PE cycle limit

Effect of type of flash (SLC versus MLC)

Common expectation:
Lower error rates under SLC ($$$) than MLC

Effect of type of flash (SLC versus MLC)



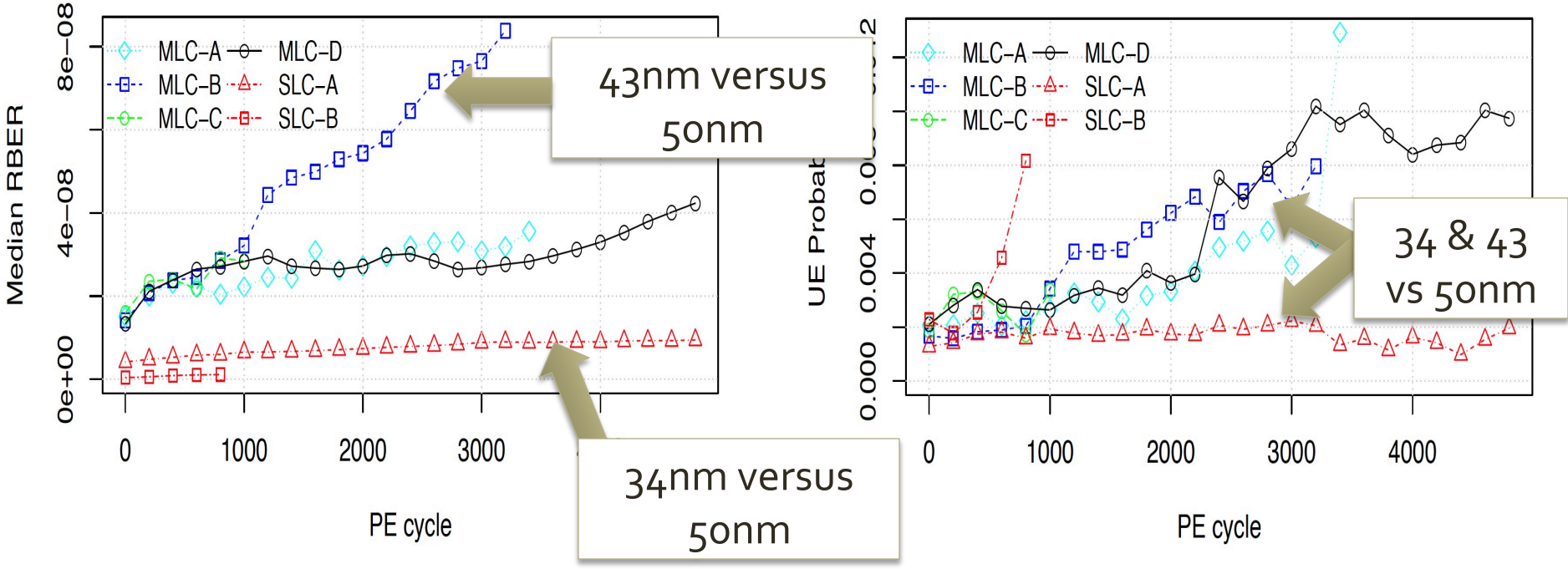


- RBER is lower for SLC drives than MLC drives
- Uncorrectable errors are not consistently lower for SLC drives
- SLC drives don't have lower rate of repairs or replacement

Effect of lithography

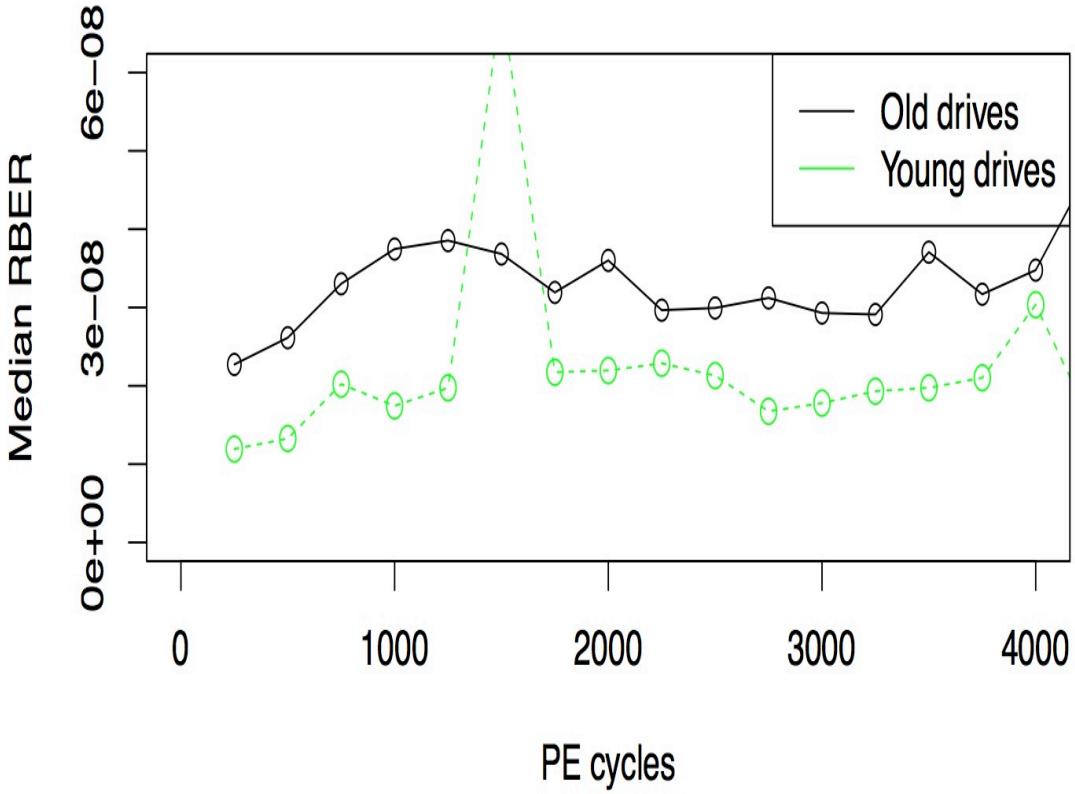
Common expectation:
Higher error rates for smaller feature size

Effect of lithography



- Smaller lithography => higher RBER
- Lithography has no clear impact on uncorrectable errors

Effect of age (time in production)?

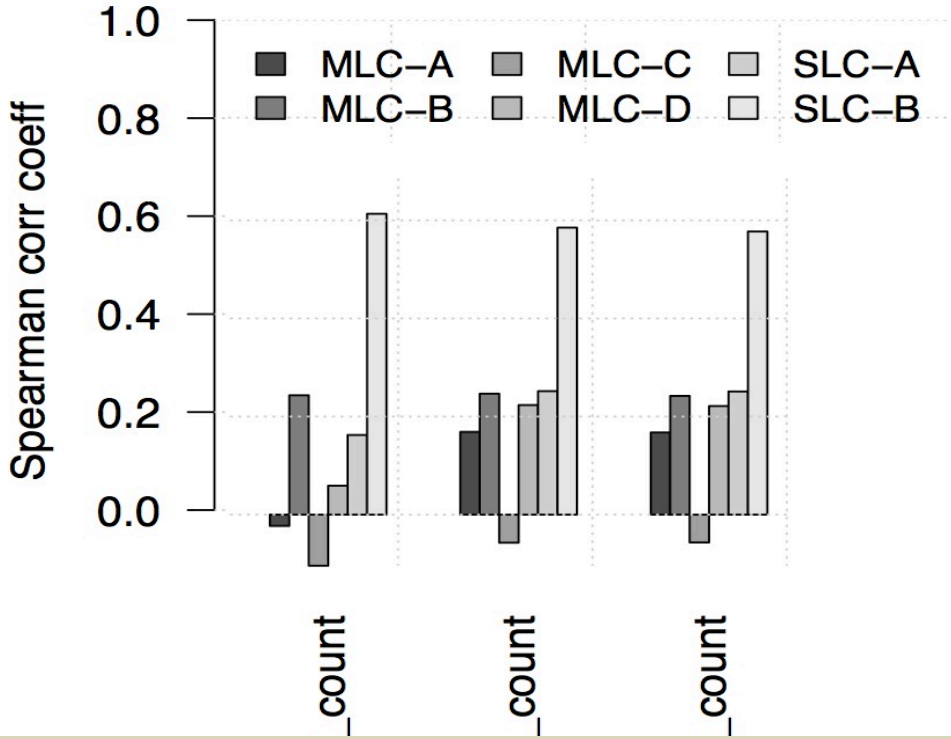


- **Age has an effect beyond PE-cycle induced wear-out**

Effect of workload?

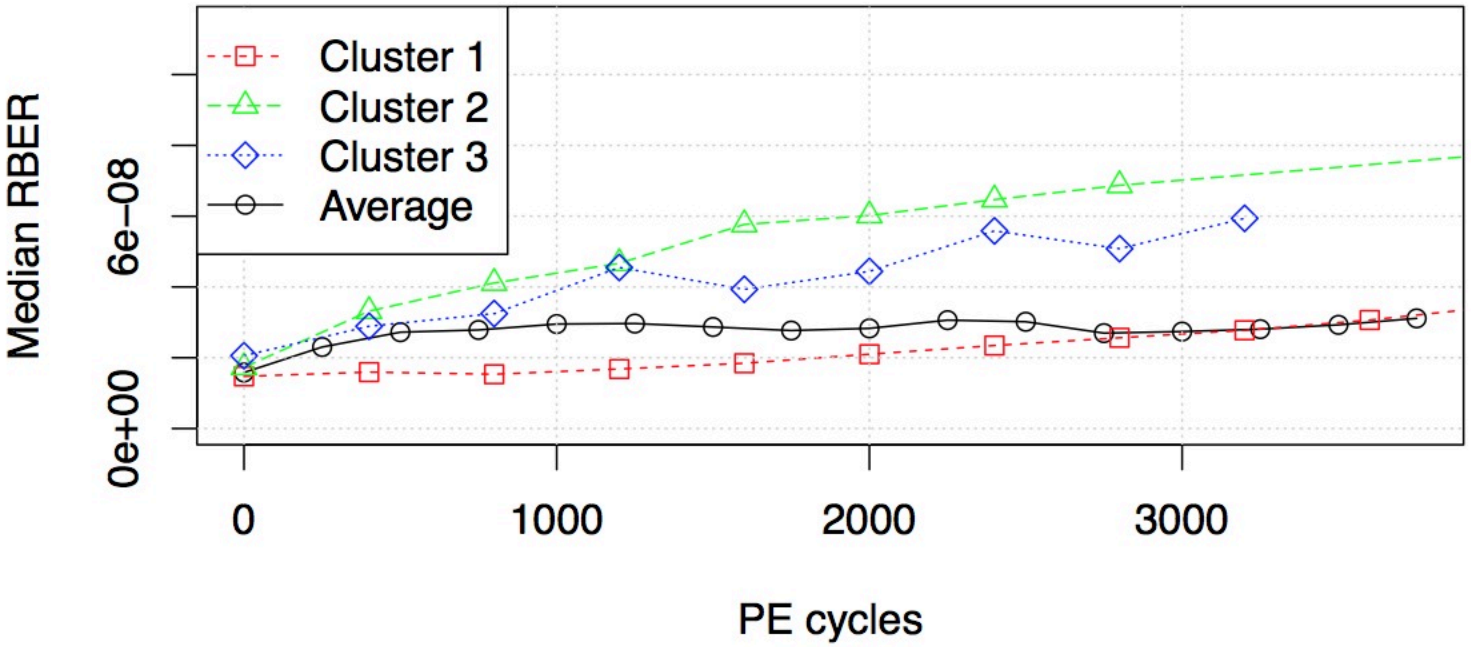
- Lab studies demonstrate workload induced error modes
  - Read disturb errors
  - Program disturb errors
  - Incomplete erase operations

How does workload affect error rates?



- Reads do affect RBER (even after controlling for PE cycles)
  - Erases and writes don't
  - Effects model dependent
- Workload does not affect uncorrectable errors
  - UBER is not a meaningful metric

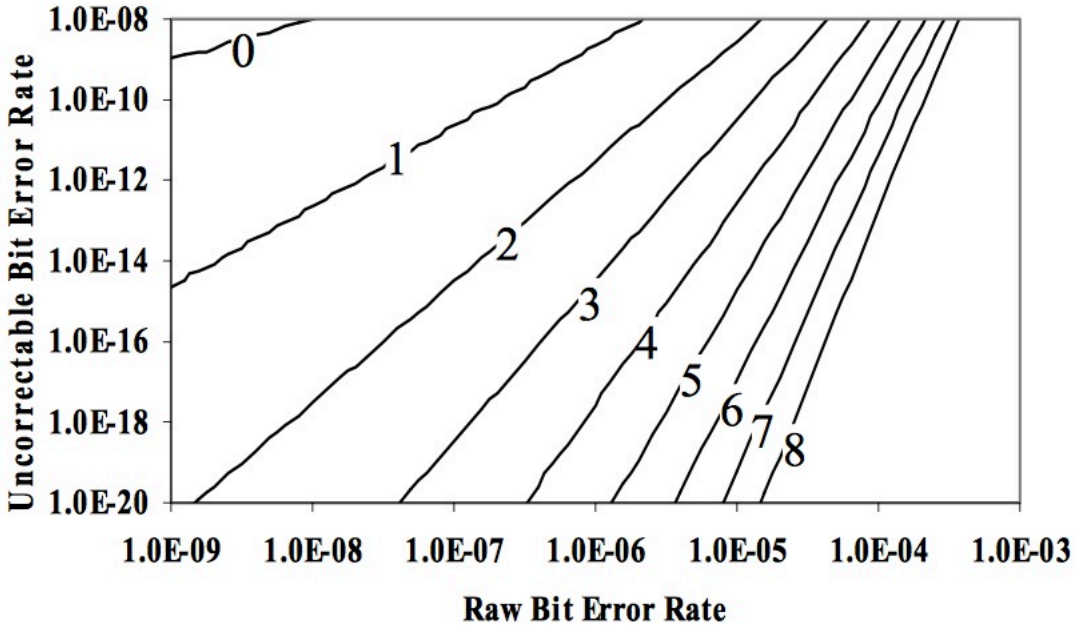
Other factors



- Different RBER for same model in different clusters
- Other factors at play …

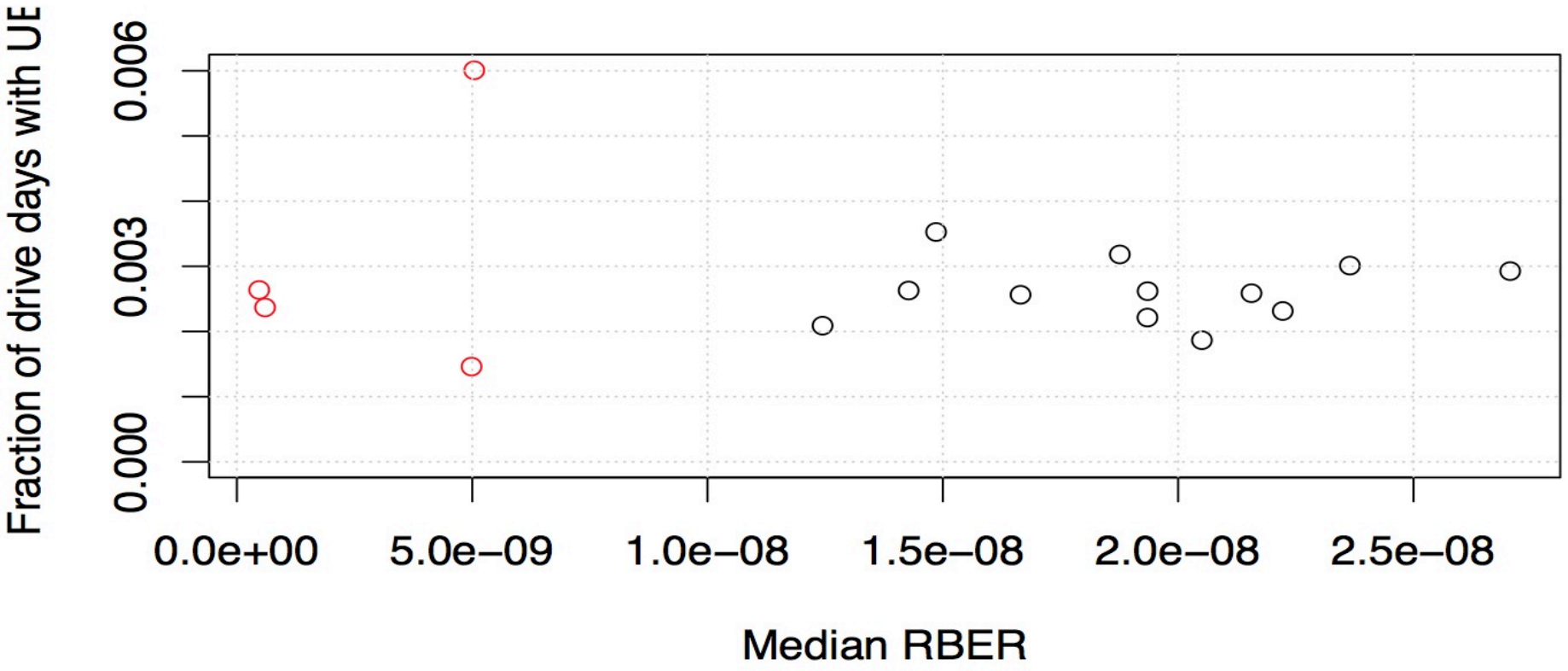
RBER and overall reliability

- The main purpose of RBER is as a metric for overall drive reliability
- Allows for projections on uncorrectable errors



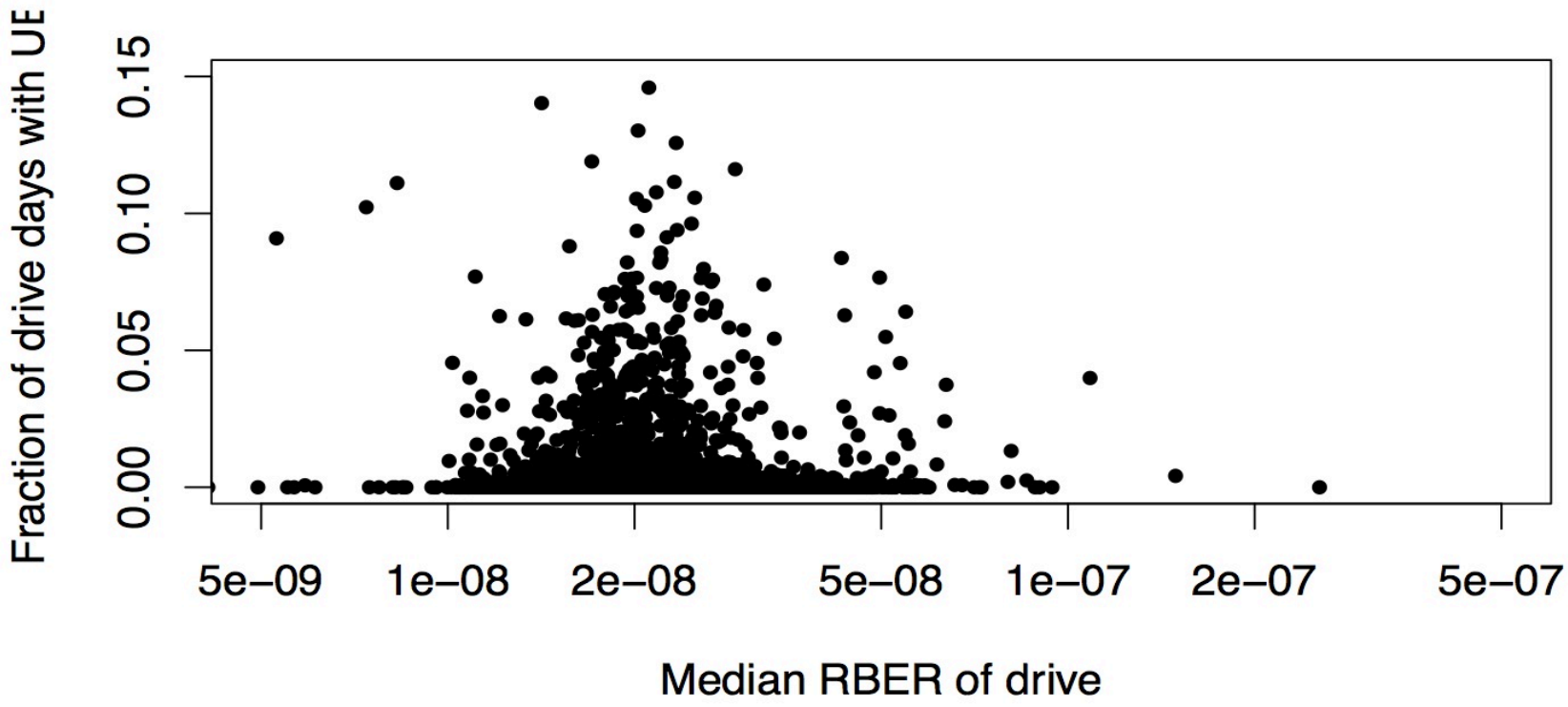
[Mielke2008]

RBER and uncorrectable errors



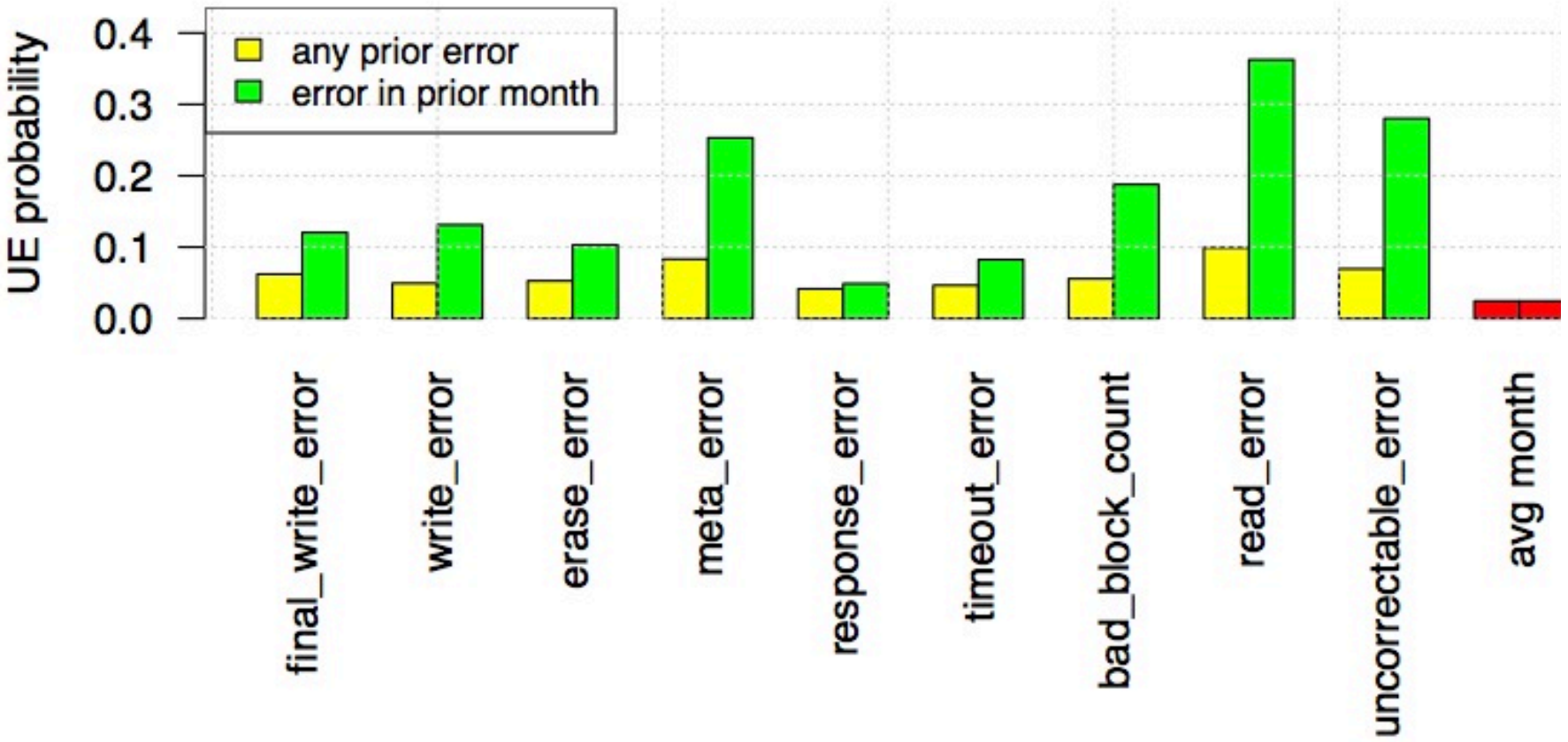
- **Drive models with higher RBER don't have higher frequency of uncorrectable errors**

RBER and uncorrectable errors



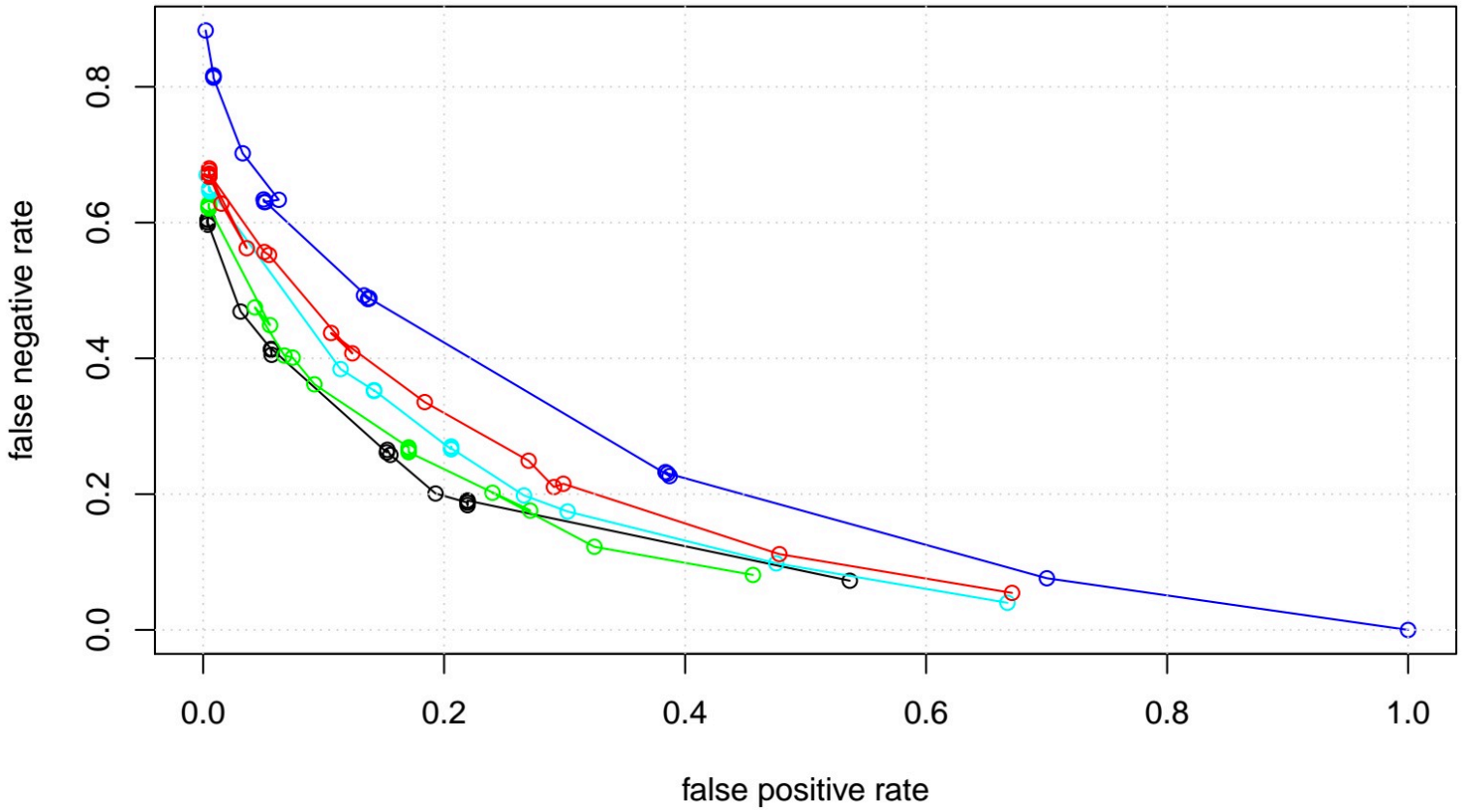
- Drives (or drive days) with higher RBER don't have higher frequency of uncorrectable errors
- RBER is not a good predictor of field reliability
- Uncorrectable errors caused by other mechanisms than corr. errors?

What is predictive of uncorrectable errors?



- Prior errors highly predictive of later uncorrectable errors
- Can we predict uncorrectable errors?

CAN WE PREDICT UNCORRECTABLE ERRORS?



- Prediction using CART models shows interesting trade-offs
  - Can catch ~30% of errors at low <0.5% false positive rate
  - Can catch 80% of errors at 20-30% false positive rate

More in the paper, that's not in the talk ….

- Comparing field RBER and prior projections based on accelerated life tests
  - Real RBER hard to predict
- Study of bad blocks & factory bad blocks
  - Vast differences between models
  - Can quickly degrade to bad chips
  - Factory-bad blocks predictive
- Study of bad chips
- Closer look at repair and replacement rates

Full paper published at Usenix FAST'2016: *"Flash Reliability in Production: The Expected and the Unexpected"* with Raghav Lagisetty and Arif Merchant.

Flash reliability – key points

- Significant rate of non-transparent errors
  - Higher than hard disk drives
  - To some degree predictable
  - Need to protect against those!
- Many aspects different from expectations
  - Linear rather than exponential increase with PE cycles
  - RBER not predictive of non-transparent errors
  - SLC not generally more reliable than MLC
- Many other results not covered in talk …