Netherlands Forensic Institute
Ministry of Security and Justice

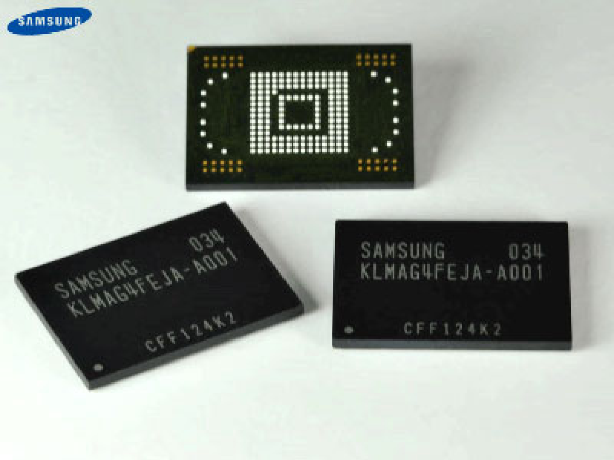
Bit-errors as a source of forensic information in NAND-flash

dr. Jan Peter van Zandwijk

j.p.van.zandwijk@nfi.minvenj.nl

Introduction (1)

NAND-flash is most popular medium for non-volatile data-storage in modern consumer electronics.



Introduction (2)

Research question:

Can bit-errors in NAND-flash memories be used for forensic purposes?

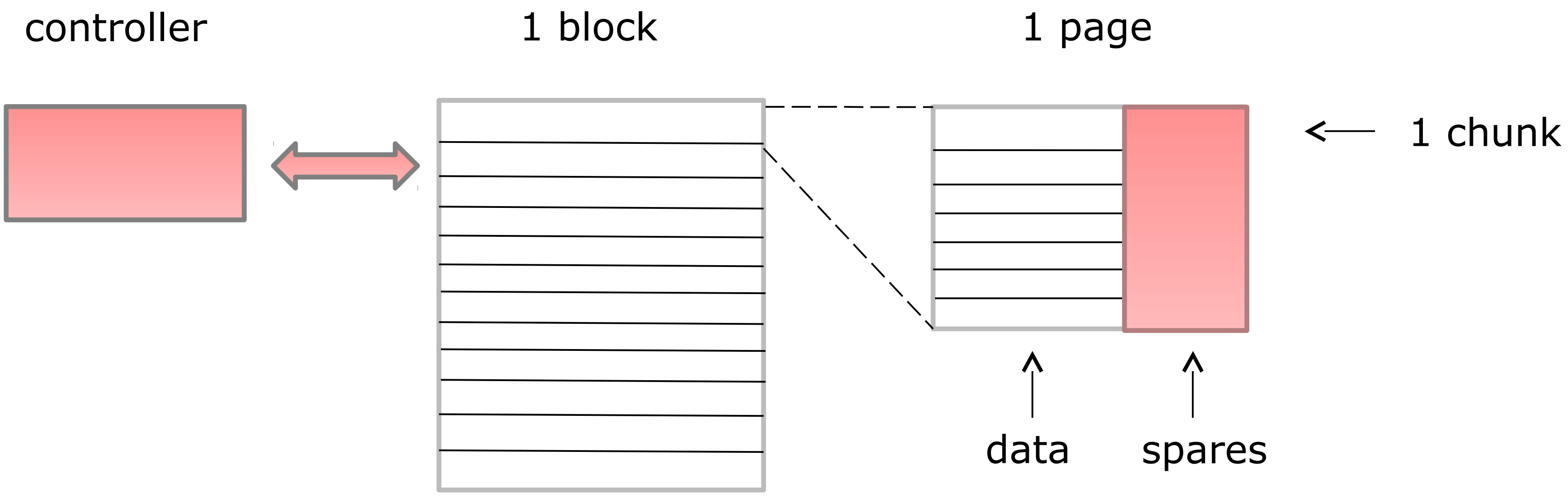
➔ NAND-flash: reliability - and error issues

➔ NAND-flash: obtaining error information

➔ Experiments and methods

NAND-flash (1)

Data is organized hierarchically in NAND-flash



NAND-flash (2)

NAND-flash contains memory cells each holding one (SLC) or more (MLC, TLC) bits of information.

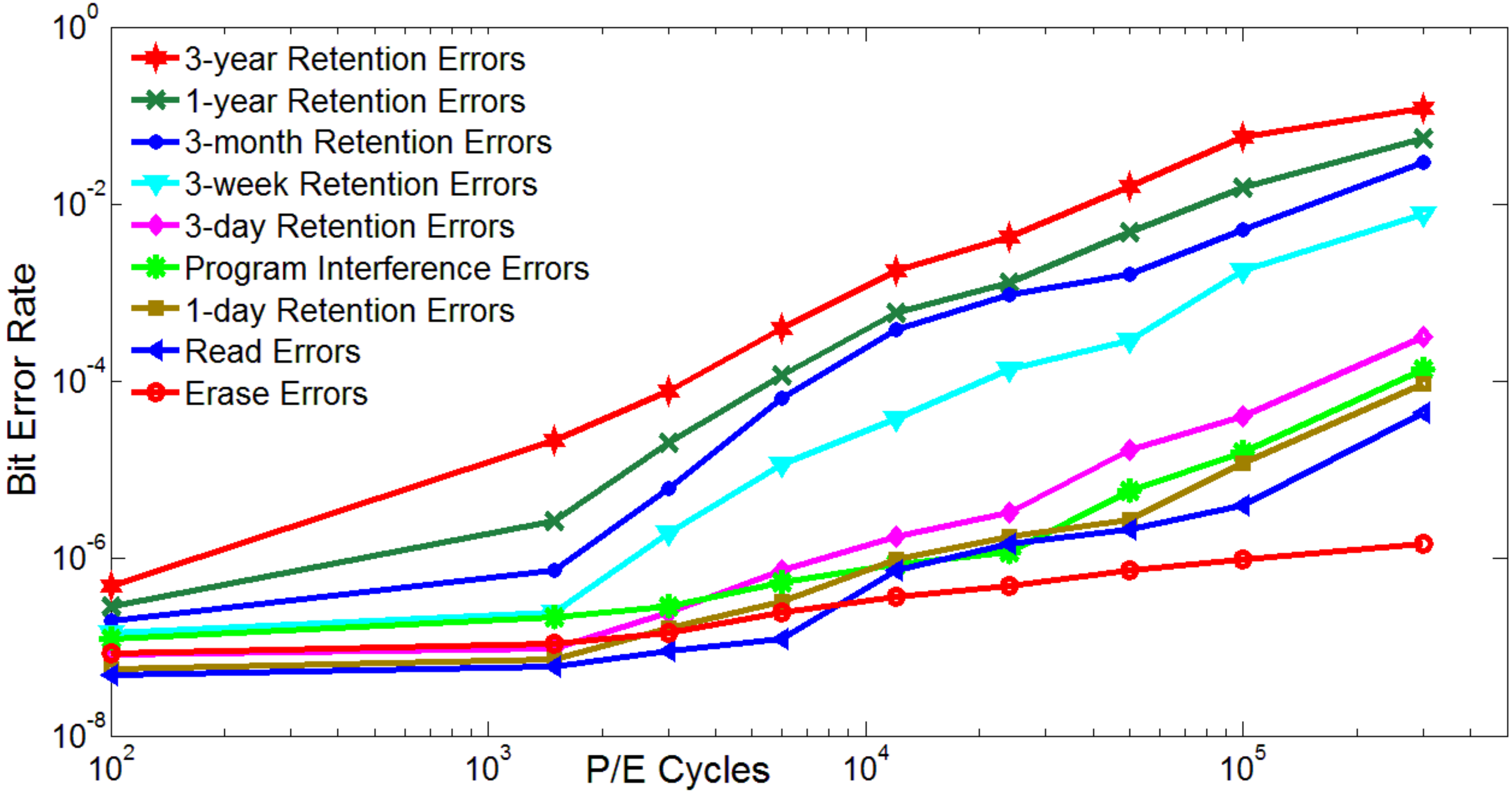
Memory cells are somewhat unreliable → content may change unintentionally.

Processes causing bit-errors include:

- Retention time
- Reading and writing of data

Rate at which error develop increases as cells deteriorate by P/E cycles.

NAND-flash (3)



From: Cai, Haratsch, Mutlu, Mai (2012)

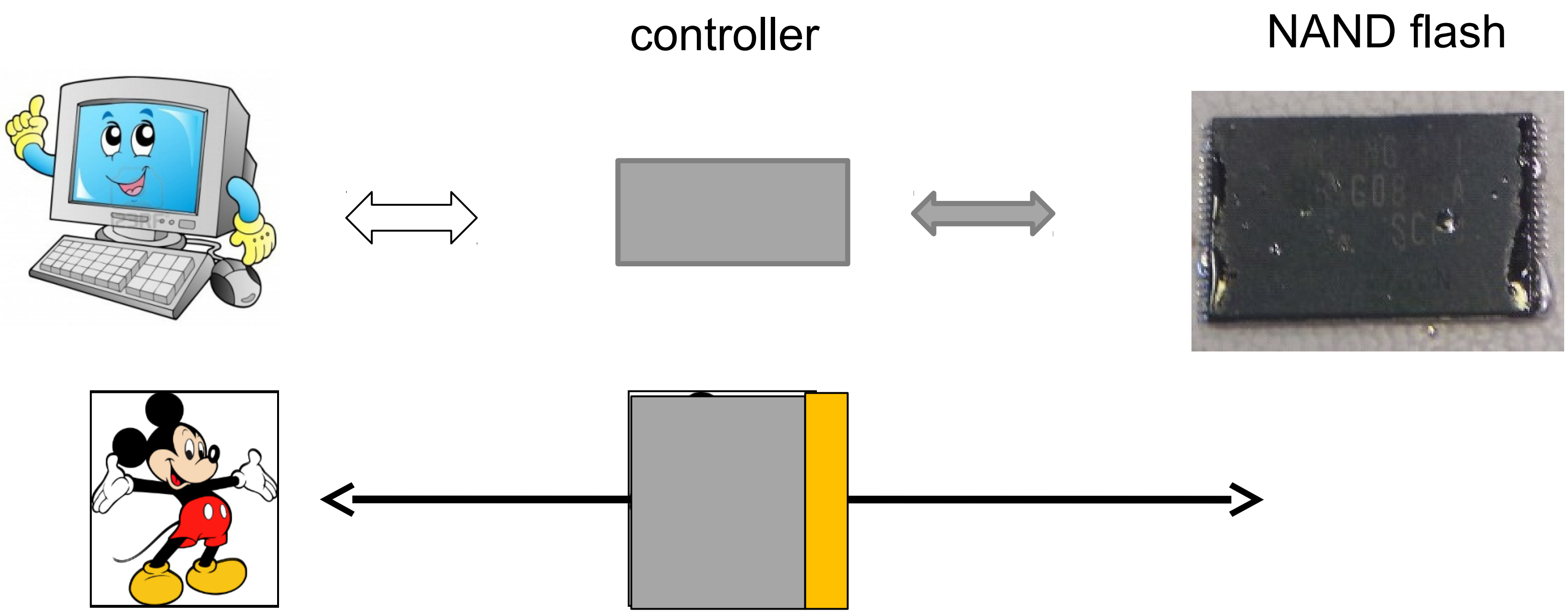
NAND-flash (4)

To improve reliability, controller uses:

- Randomisation (XOR-pattern)
- Error-correcting codes (ECC)
- Wear leveling

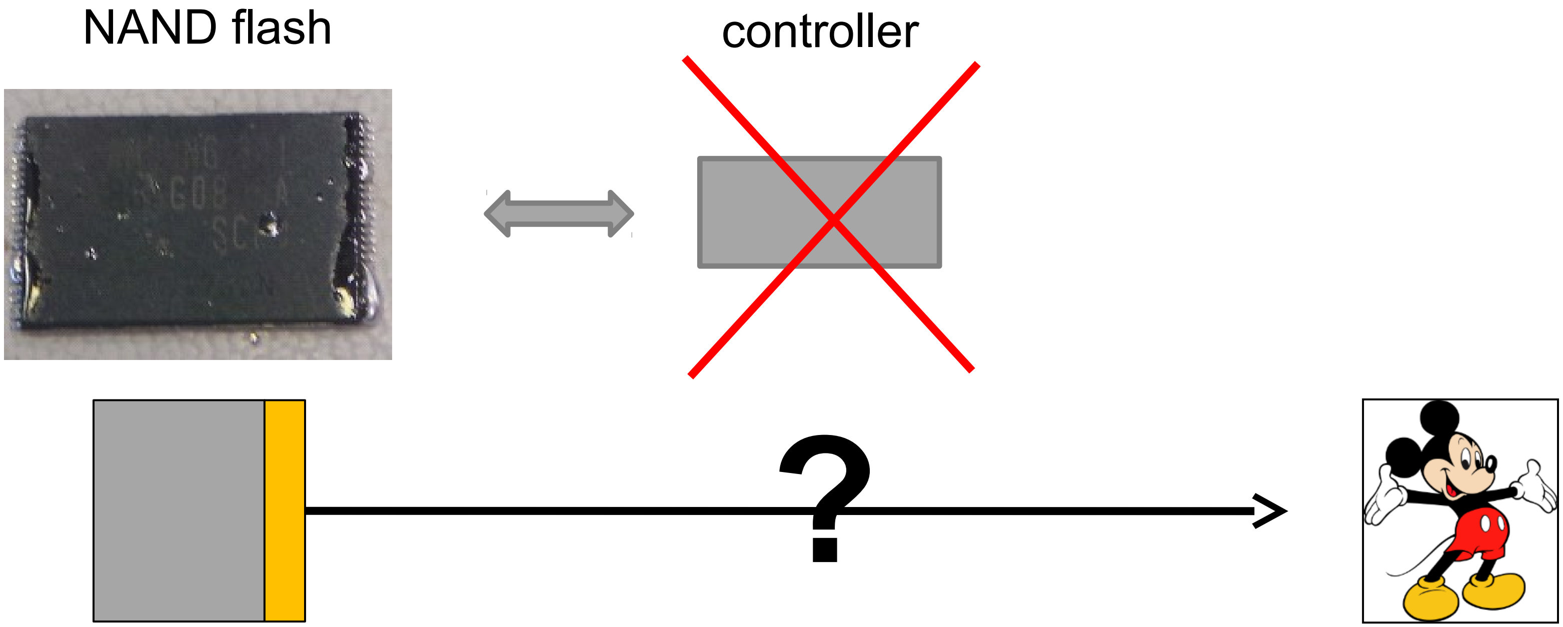
➔ Actual data stored on NAND-flash is different from data in write request from host OS.

NAND flash (5)



Typical sequence of data storage on NAND-flash: randomisation and ECC parity-bit computation

NAND flash (6)



NAND flash

controller

Question: without using the controller, can content of memory pages be exposed by removing randomisation and application of ECC parity-bits to correct bit-errors?

NAND flash (7)

1. **Reconstruction of XOR-pattern**

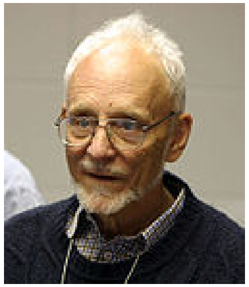
   - Assume randomisation is produced by LFSR
   - Use BM-algorithm to determine LFSR feedback taps

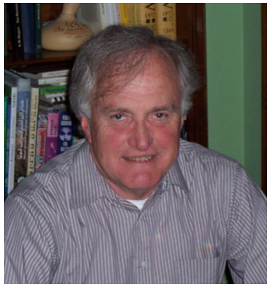
2. **ECC reconstruction**

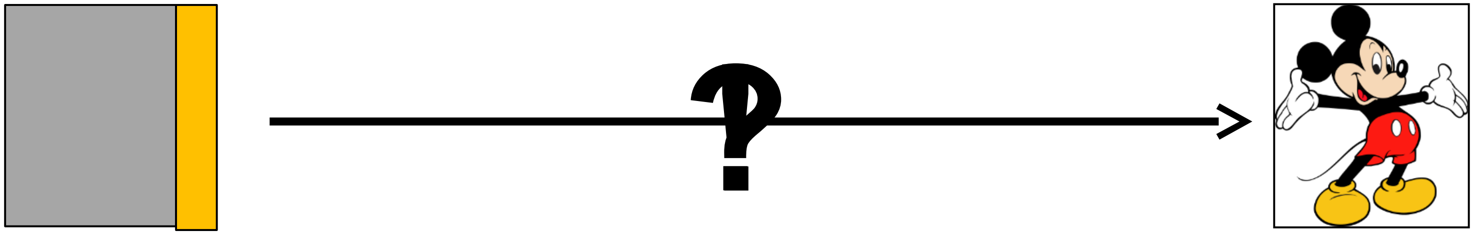
   - Assume cyclic code in use.
   - Stepwise reconstruction of code parameters
     - generator polynomial g(x)
     - length n, dimension k, correctable errors t
     - BCH parameters

NAND flash (8)

NAND flash







Reference:
J.P. van Zandwijk: A mathematical approach tot NAND flash-memory descrambling and decoding. Digital Investigation 12 (2015) 41-52
Available after the talk …

NAND flash (9)

Research question:

Can bit-errors in NAND-flash memories be used for forensic purposes?

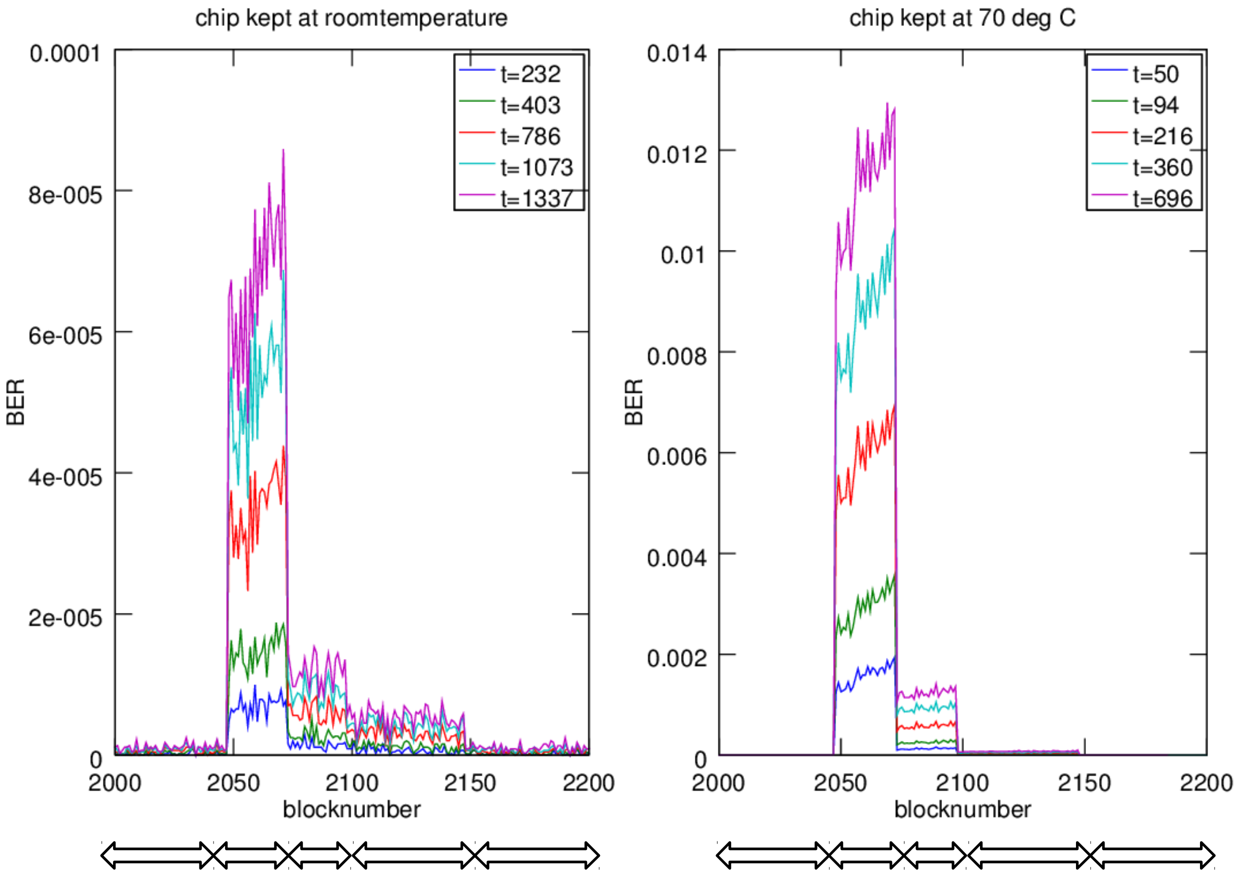
- ➔ Bit-errors are related to forensically interesting information, such as e.g. retention time
- ➔ Is it detectable on forensically relevant timescales?
- ➔ How well can it be related to activities performed on the NAND-flash?

Experiments (1)

Raw NAND-flash experiments

- Check for detectable retention bit-errors on forensically relevant timescales when used within factory specifications in terms of P/E cycles
- Known random data directly written to and read from memory, no ECC
- Used high temperature baking to simulate longer retention times
- Errors computed by comparison of data from chip with original data

Experiments (3)


chip kept at roomtemperature

chip kept at 70 deg C

# P/E cycles: **0** **2500** **1200** **500** **0** **0** **2500** **1200** **500** **0**

Experiments (4)

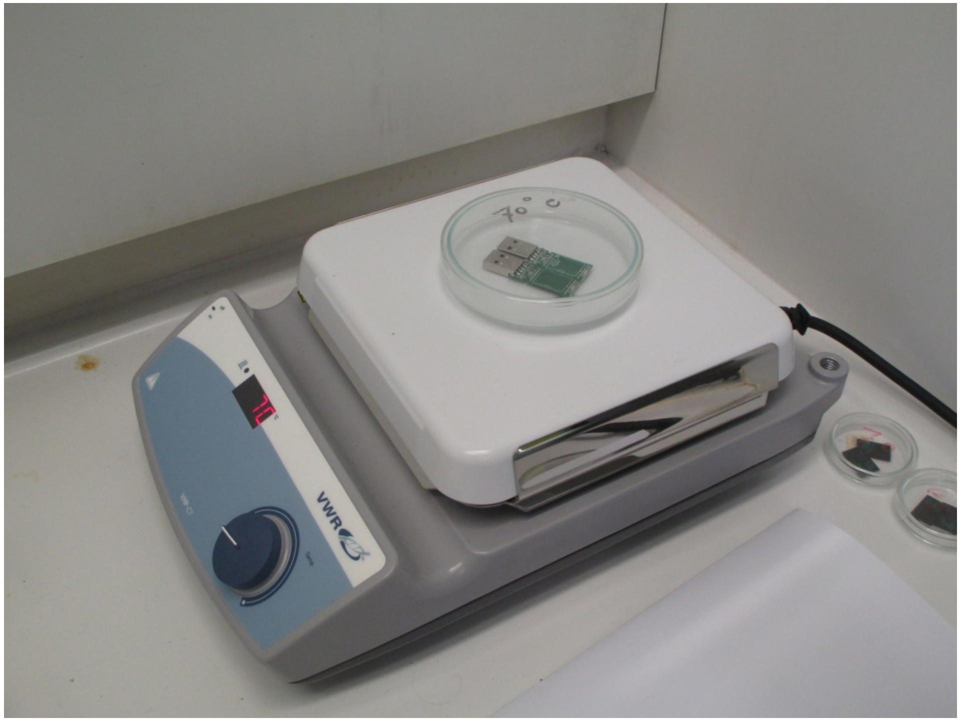
Raw-NAND conclusions

- Detectable difference in retention bit-errors over periods of weeks-months
- Number of errors depends on storage temperature
- Seizable effect of number of P/E cycles on rate of bit-error development, even for moderate number of P/E cycles

Experimenten (5)

USB thumb-drive experiments

- Performed user-activities on USB thumb-drives
- NAND-flash desoldered and read
- Processed raw dump to get bit-error stats.
- Checked for relation between actions and bit-errors stats

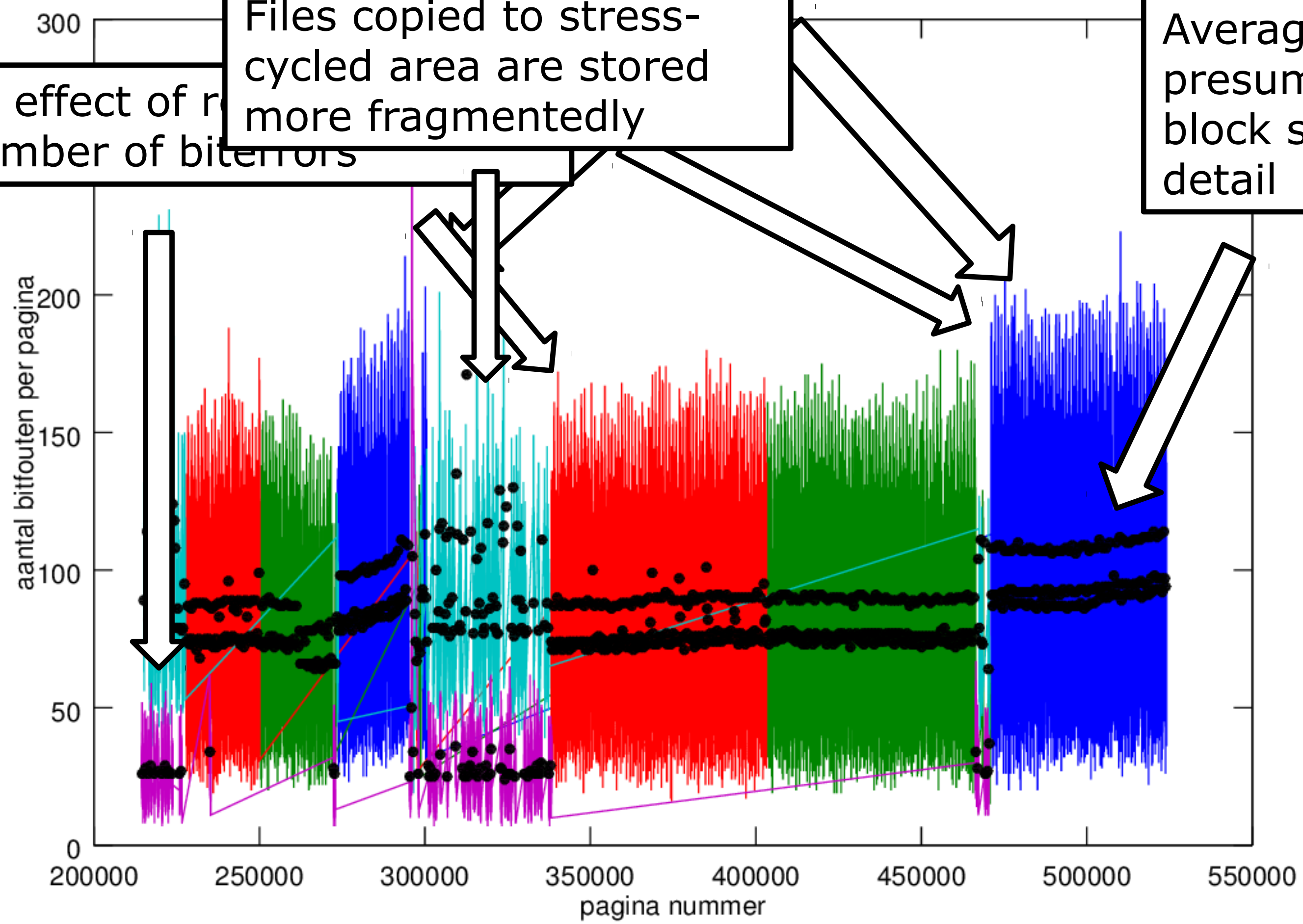


Experiments (6)

USB thumb-drive experiments

- Copied files with random data onto drives at different times through the controller
- Drives stored at either RT or 70^{0}C in-between
- Two protocols:
  - data retention
  - drives partially stress-cycled by repeatedly copying and deleting data
- Offline data analysis:
  - decoding of data yields bit-error statistics for pages
  - descrambled pages tied to files using hashes

Experiments (7)



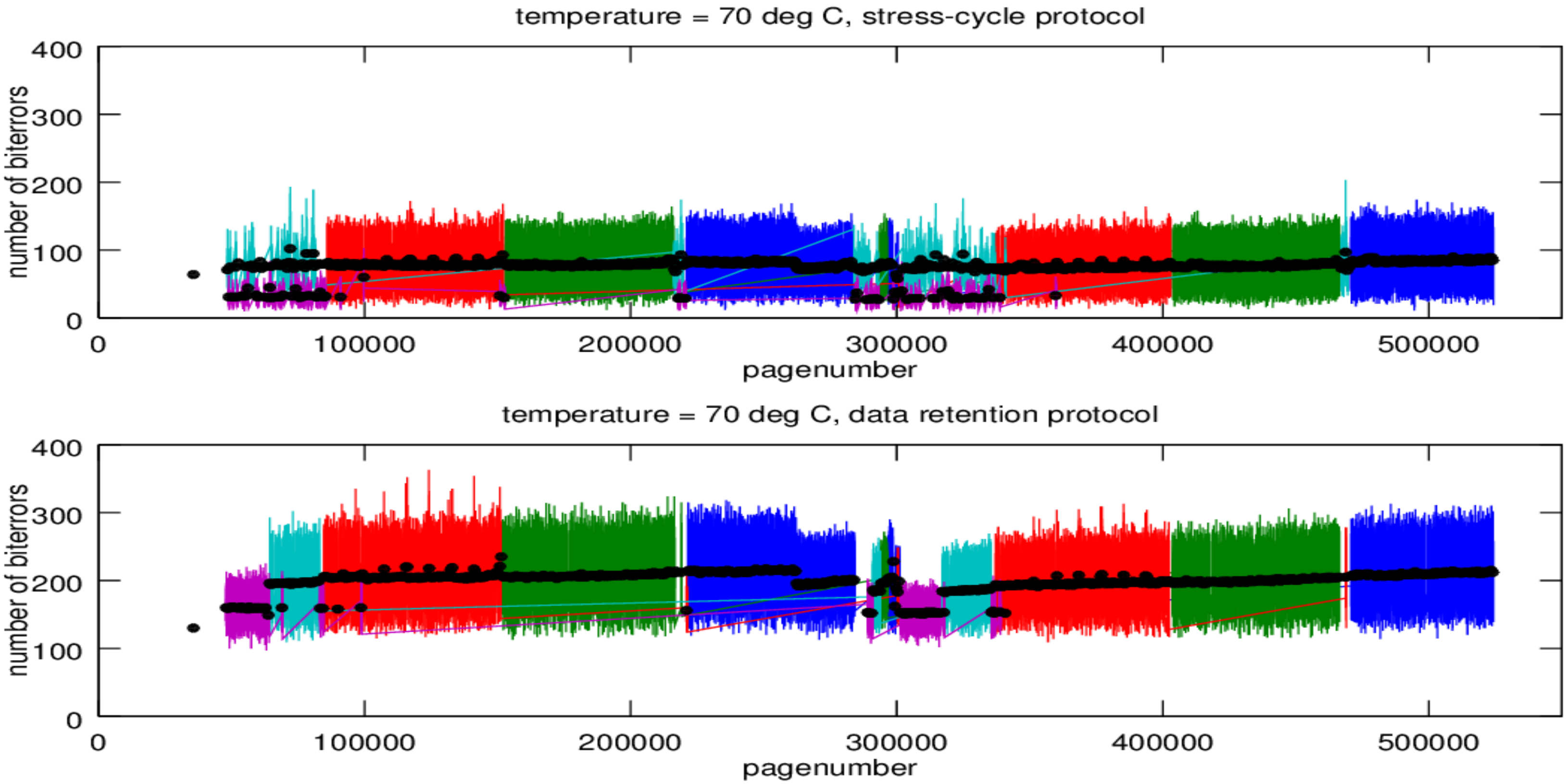
Spread of memory pages due to wear levelling

Files copied to stress-cycled area are stored more fragmentedly

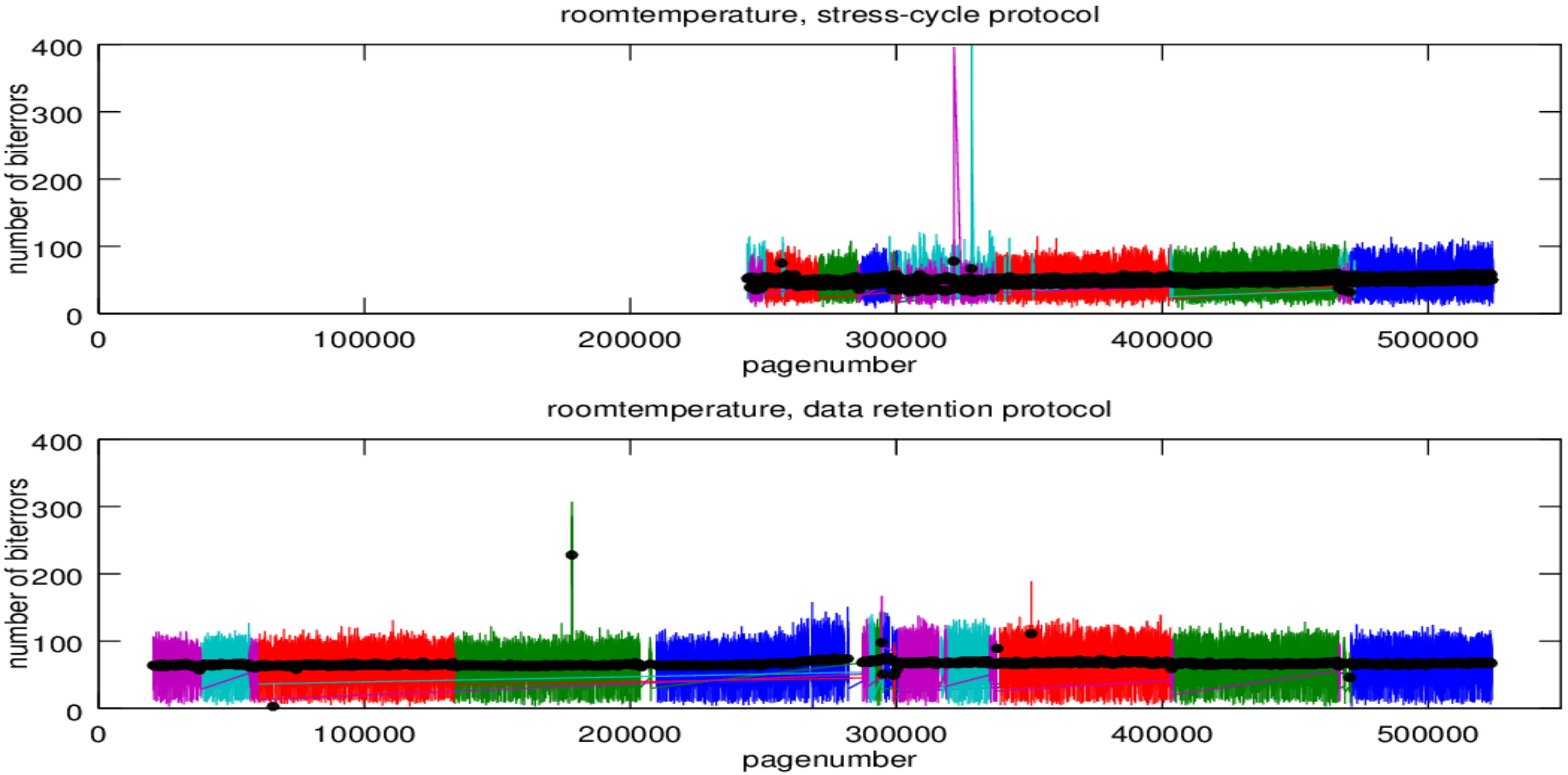
Some effect of r
on number of biterrors

Averaging over presumed erase block shows more detail

Experiments (8)



Experiments (9)

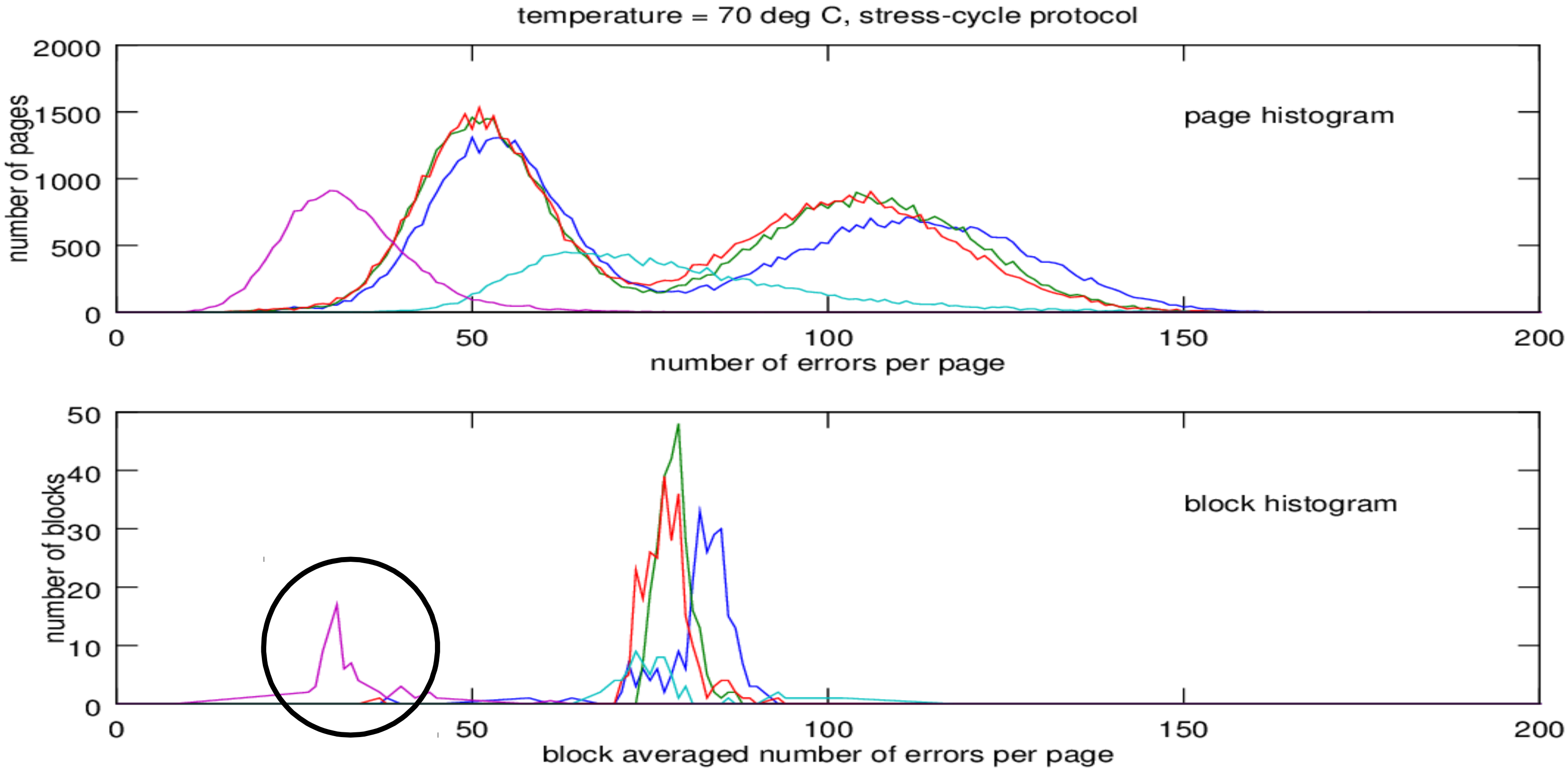


Experiments (10)

USB thumb-drive conclusions

- Area used for data-storage varies between drives
- Baseline error-level varies for drives stored at the same temperature
- For drives stored at higher temperature, some relation between number of bit-errors and retention time
- Copying data onto stress-cycled areas leads to fragmented data storage

Experiments (11)



Number of block errors can be specific for files

Conclusions (1)

Low-level acquisition and off-line analysis provides access to an independant side-channel of a normally functioning device, not accessible otherwise.

Possible use of side-channel:

- Provide an independant time side-channel for NAND-flash based devices, such as e.g. SSDs
- Perform (relative) dating of memory pages in NAND flash.
- Use as a means of grouping memory pages to aid smart carving and file reconstruction.

Conclusions (2)

Caveats:

- Many factors affecting bit-error statistics, currently only limited initial research
- In real life, different factors may occur jointly or concurrently
- Might be difficult to separate contributions of e.g. retention time and P/E cycles
  - can ‘old’ data on a fresh page be distinguished from ‘new’ data on a worn out page?