



***MicroSys***

**Erhöhung der Systemsicherheit  
durch die Wahl ECC geeigneter  
Prozessoren und Speicher**


**ECC 2017**

**Kei Thomsen**

# Hintergrundinformation zu immer kleineren Chipstrukturen

- Chip Struktur bei DDR4 RAM aktuell bei 10nm
- 1.2V Spannung
  - reduzierte Leistung
  - höhere Geschwindigkeit
- Gefahr von Soft-Error-Störungen durch Single Event Upsets (SEU)
  - Höhen- und radioaktive Strahlung
  - Fehlerrate steigt mit der Höhe über Grund
  - Energiereiche Strahlung von Elektromotoren und statischen Feldern
- Ungenaue Timing Einstellungen beim RAM Controller
  - RAM Funktioniert im Labor bei „Raumtemperatur“
  - Was ist bei  $-30^{\circ}$  oder  $+80^{\circ}$ ?
- Abhilfe: Fehlererkennung und Behebung

# Erhöhung der Zuverlässigkeit durch ECC Speicher

- ECC = Error Correction Code.
- Für 1 Bit Error Correction und 2 Bit Error Detection
  - 32 Bit Speicher: 7 Bit extra
  - 64 Bit Speicher: 8 Bit extra
- ECC fähiger Memory Controller im SOC (System On Chip)
  - Einige wenige ARM CPUs 
  - X86 fast nur bei Server-Chipset
  - Fast alle PowerPC CPUs und LayerScape (ARM)
- Zu erkennen an 5 oder 9 RAMs auf dem Board
- Controller erzeugt einen Interrupt bei auftretendem Fehler
- Betriebssystem sollte das Ereignis entsprechend behandeln
- Test / Simulation durch Error Injection Unit



# NAND Flashes und deren Sicherung gegen Datenverlust und bevorstehende Katastrophen

## Typen von NAND Flashes:

- SLC = Single Level Cells. 1 NAND Zelle = 2 Zustände = 1 Bit
- MLC = Multi Level Cells. 1 NAND Zelle = 4 Zustände = 2 Bit (SCHWARZE MAGIE)
- TLC = Three Level Cells. 1 NAND Zelle = 8 Zustände = 3 Bit (Alien Technologie)
  - <http://www.speedguide.net/faq/slc-mlc-or-tlc-nand-for-solid-state-drives-406>
- MLC/TLC: sehr aufwendige Korrekturverfahren im Chip notwendig.
- Pro 512 Byte sind typ. 16 Byte OutOfBand (OOB) Daten abgelegt.
- 50nm SLC nutzt 3 Byte für ECC (1 Bit ECC) bei 100K Zyklen oder 1 Jahr Datensicherheit.
- 30-40nm TLC benötigt 8+ Bit ECC für 5K Zyklen oder 1 Jahr.
- TLC aktuell bei 12nm...

Western Digital White Paper: NAND Evolution and its Effects on Solid State Drive (SSD) Useable Life

# NAND Flashes und deren Sicherung gegen Datenverlust

Schreib-Lösch-Zyklen pro Zellblock, danach nicht mehr nutzbar

- SLC ~ 100.000 - 1.000.000
- MLC ~ 3.000 - 10.000
- TLC ~ 1.000 - 3.000

## Wear-Leveling notwendig

- Verteilt die Schreib Zugriffe über den NAND Flash Bereich
- Damit jeder Sektor gleich häufig beschrieben wird
- Sektor 0 ist damit mal in Block 0, 1527, ..., 42560, <wo auch immer>
- Betriebssystem Level: UBIFS, JFFS2, YAFFS2 etc.
- Chip Level: SD,  $\mu$ SD, CF, USB-Stick, SSD etc

# NAND Flashes und bevorstehende Katastrophen

## BitFlips nach einigen Monaten

- 1-2 mal pro Tag wird der NAND (SLC) Bereich gelesen (Linux Image laden im U-Boot) oder das Board liegt unbenutzt im Schrank.
- Nach 6-12 Monaten treten einzelne BitFlips in Sektoren auf.
- Das lässt sich erst mal nicht erkennen, da der Controller durch ECC die Daten selbstständig korrigiert hat.
- Nach 12-18 Monaten treten doppelte BitFlips in Sektoren auf.
- Nun lässt sich das Linux Image gar nicht mehr lesen
- **Wenn so etwas beim Anwender passiert, ist das Geschrei groß.**



# NAND Flashes und Verhinderung von Datenverlust

Wie lassen sich BitFlips korrigieren/verhindern?

- Methode 1, Korrektur (alle 1-2 Monate)
  - Regelmäßig alle Sektoren lesen und den NAND Flash Controller auslesen, ob ein Single Bit Flip vorlag, den er korrigieren konnte.
  - Diesen Sektor neu schreiben, mit den (noch) korrekten Daten.
  - Vergleich ob jetzt okay, sonst Sektor als Defekt markieren und ummappen.
- Methode 2, Vorbeugung (alle 1-2 Monate)
  - Regelmäßiges Scrubbing. Dabei werden alle Sektoren eines Bereichs gelesen und neu geschrieben, bzw. umgemappt.
- Bei SD, µSD, CF, SSD, USB-Stick wird (sollte) es vom Controller in dem Device organisiert.
- Bei eMMC müssen dazu regelmäßig alle Sektoren gelesen werden.
- Nur was ist mit einem Backup auf SD oder CF, das für 3 Jahre im Schrank liegt?

# NAND Flash Horror Story

- S\*\*Disk CF Karte > 2GB (Consumer Karte)
- Partial Wear Leveling... nur im FAT32 Bereich (e.g. erste 4MB)
- Mit FAT32 okay, z.B. für DSLR Kameras.
- Mit OS-9, QNX, Linux (EXT2/3/4) Filesystem
  - Nach ein paar 1000 Schreibzugriffen auf das Filesystem ist die Karte defekt.
  - Diese Filesysteme legen die Tabellen/Einträge mehrfach ab, damit Defekte in einzelnen Sektoren nicht das Filesystem zerstören.
  - Aber genau das zerstört die NAND Flashes und damit das Device.
- Auswahlkriterien für Industrietaugliche CF, SD und  $\mu$ SD Karten
  - SLC für höchste Qualität, MLC wenn akzeptabel, TLC no-go
  - Full, Static oder Dynamic (weil klingt besser als Static) Wear-Leveling.
  - Auf keinen Fall **Partial Wear Leveling**



# SSD Festplatten

- SSD Platten basieren auf NAND Flash. SLC, MLC & TLC.
- Je lauter die Werbung, desto mehr wird der Billigkram hochgespielt:
  - „Bahnbrechende 3D-NAND-Technik“ bezeichnet meist TLC (ev. MLC), bei denen mehrere Chips/Dies übereinander angeordnet sind.
- SSD mit SLC muss man speziell suchen und sind recht teuer.
- Umfrage bei  Mitgliedern:
  - „wir haben in 10 Rechnern seit 1 ½ Jahren SSD Platten im Entwicklungseinsatz und innerhalb von 2 Wochen sind 3 SSD Platten so ausgefallen, dass gar nicht mehr ging! Noch 1 Woche vorher zeigten die SMART Daten keine Probleme an.“
  - „SSD... für Entwicklungsrechner ja, für industriellen Einsatz niemals!“
- Peta-Byte Club. Test von 250GB SSD Platten, bei dem versucht wird, 1 Peta Byte (1024 Tera Byte) zu schreiben.
  - Nur 3 von 6 Platten haben es über 700 Terabyte geschafft.

[techreport.com/review/26523/the-ssd-endurance-experiment-casualties-on-the-way-to-a-petabyte](http://techreport.com/review/26523/the-ssd-endurance-experiment-casualties-on-the-way-to-a-petabyte)

# Zusammenfassung

- Sicherungsmaßnahmen gegen die kleiner werdenden Strukturen und den eingesetzten Alien Technologien.
- ECC Speicher und SOCs mit entsprechenden Controllern erhöhen die Systemsicherheit und Stabilität ganz massiv.
- NAND Flashes für den industriellen Einsatz bitte nur mit SLC Technologie, alles andere zu gefährlich und nicht für den 10 Jahre Dauereinsatz.
  
- **Fragen?**
- Gerne auch am Stand von Elma/MicroSys
  
- Weiterer Vortrag im Stream3 16:30 :  
**Hard- und Softwareaspekte zur Optimierung von embedded Systemdesigns**