

# DRAM Speicherfehler im Betriebssystem Problem und Gegenmaßnahmen

Vortrag im Rahmen des Masterseminars „Fehlertolerante Systeme“

Robert de Temple

Erlangen, 11.11.2014



# Inhalte



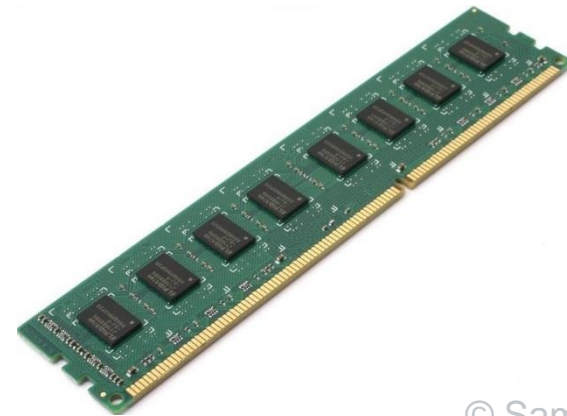
# Motivation

## Ein PC heute



© Intel

- 4 Kerne + GPU
- ~ 2 Mrd. Transistoren



© Samsung

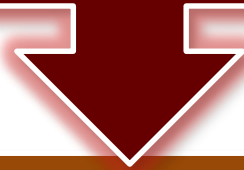
- 16 GiB
- > 17,2 Mrd. Transistoren  
+ 17,2 Mrd. Kondensatoren

# Motivation

- Speicherfehler gehören zu den drei häufigsten Fehlern in modernen Rechensystemen
- Speicherfehler haben eine besonders hohe Chance für SDC (silent data corruption)
- Das Betriebssystem ist besonders anfällig
  - ▶ hier oft Systemversagen

# Speicherfehler

Heute + in Zukunft



# Gegenmaßnahmen

Stand der Technik



# Neue Ansätze

Software + Hardware

# Speicherfehler

Heute + in Zukunft

**Heute**

- DRAM -

**Zukunft**

- PCM-RAM -



**Gegenmaßnahmen**



**Neue Ansätze**

# DRAM – Einführung

## Ladung = Information

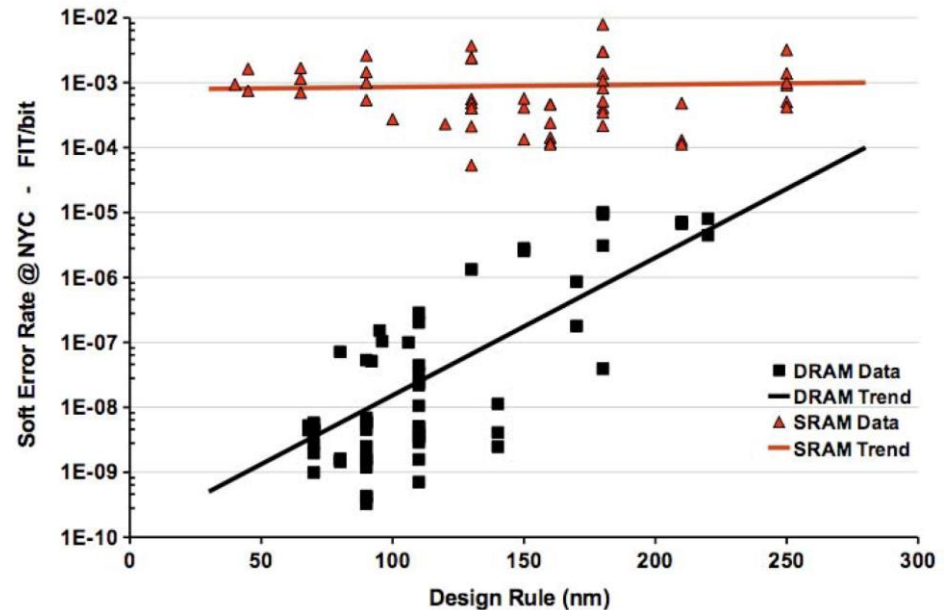
(Ziegler et al. 1979)

### Fehlerquellen

- Alphastrahlung
- Neutronenstrahlung

### Fehlermodell

- Transiente Fehler



Slayman et al., CSME 11'  
„Soft errors in modern memory technology“

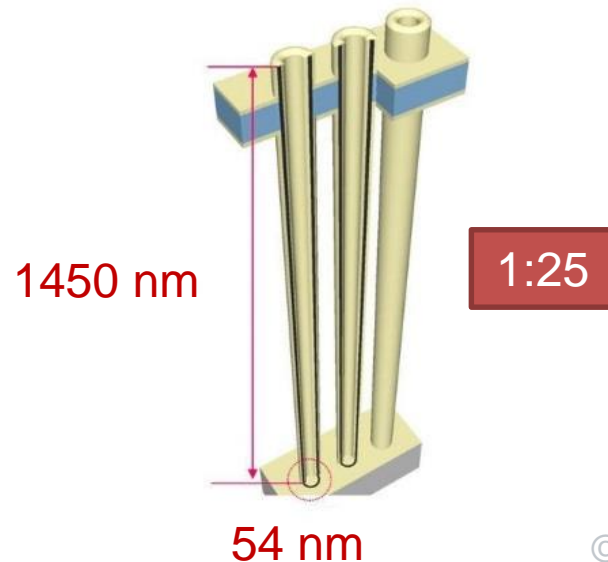
- FIT/Bit nimmt mit Fertigungsgröße ab
- Wahrscheinlichkeit für Multi-Bit-Fehler nimmt dabei zu
- FIT/Chip nimmt mit Fertigungsgröße auch ab

# DRAM – Einführung

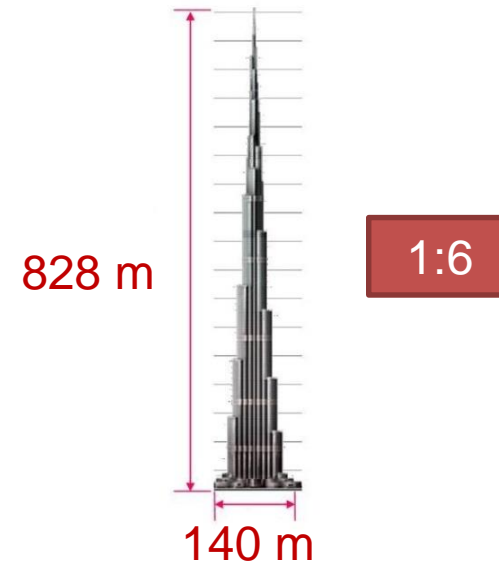
## Gründe

- Kapazität des Kondensators skaliert kaum
  - ▶  $Q_{\text{krit}}$  fast konstant geblieben
- Kondensatoren immer tiefer, dabei dichter gepackt
- Steuerungselektronik ist kleiner, dichter

3x nm DRAM



Burj Khalifa



© weSRCH.com



# DRAM – Schroeder et al.

**Schroeder et al. (2009)**

**University of Toronto**

- Untersuchung von Speicherfehlern in Großrechnern (BlueGene, Google-Server, SciNet)

## Ergebnisse

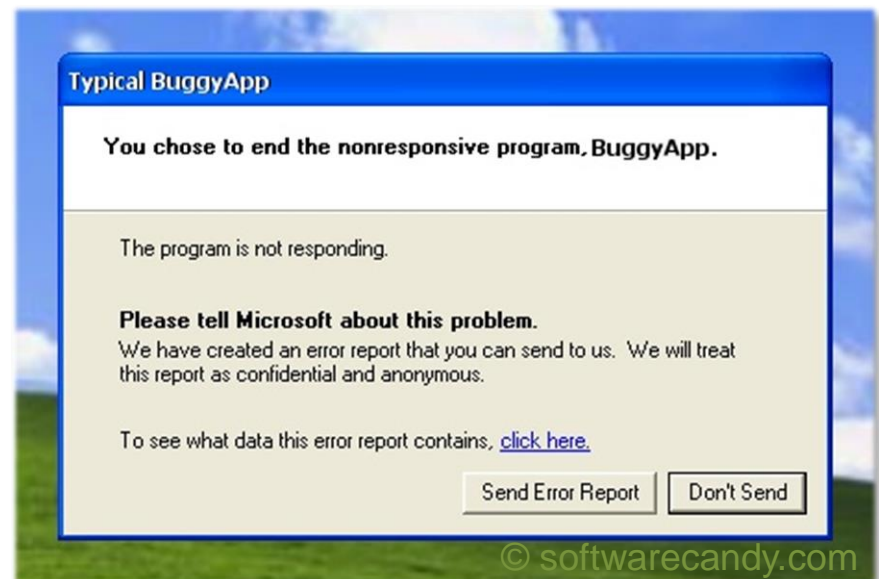
- $\emptyset$  8% aller RAM-Riegel / Jahr von Fehlern betroffen
- 60 – 80% eindeutig permanente Defekte
- Meist komplexes, sporadisches Verhalten
  - ▶ schwer zu reproduzieren
- Fehler ungleichmäßig im Speicher verteilt
  - ▶ Vom BS genutzter Speicher besonders häufig betroffen

# DRAM – Nightingale et al.

Nightingale et al. (2011)

Microsoft Research

- **Windows Customer Experience Improvement Program (CEIP)**
  - ▶ Ergebnisse für Consumer PCs
- Beobachtungszeitraum 8 Monate (Jahr 2008)
- Ca. 950.000 PCs
- CPU, RAM und HDD-Fehler



# DRAM – Microsoft Research

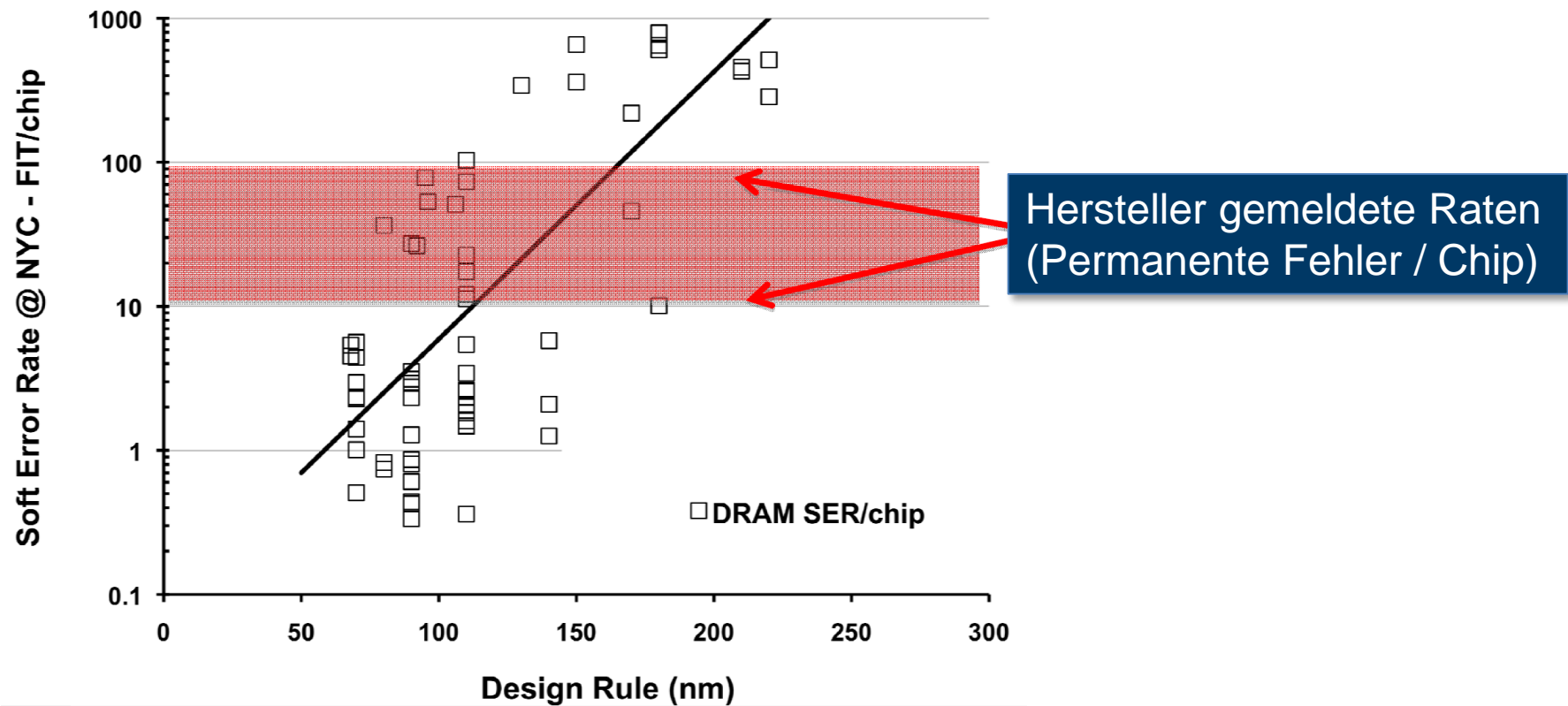
## Einschränkungen

- Ausschließlich 30 MB Kernel Abbild des Betriebssystems (~1,5% des Ø-Speichers) ► hier Systemabsturz
- Ausschließlich 1-Bit Fehler

## Ergebnisse

- Ergebnisse von Schroeder für „normale“ PCs größtenteils bestätigt
- Fehlerraten wieder viel höher als erwartet
- Ca. 15% - 19% aller Fehler traten wiederholt auf  
► davon 79% an der exakt gleichen Speicherstelle
- **Außerdem:** Speicherfehler nicht unabhängig von CPU-Fehlern

# DRAM – Industriekennzahlen



Slayman et al., CSME 11'  
„Soft errors in modern memory technology“

# Speicherfehler

Heute + in Zukunft

**Heute**

- DRAM -

**Zukunft**

- PCM-RAM -



**Gegenmaßnahmen**



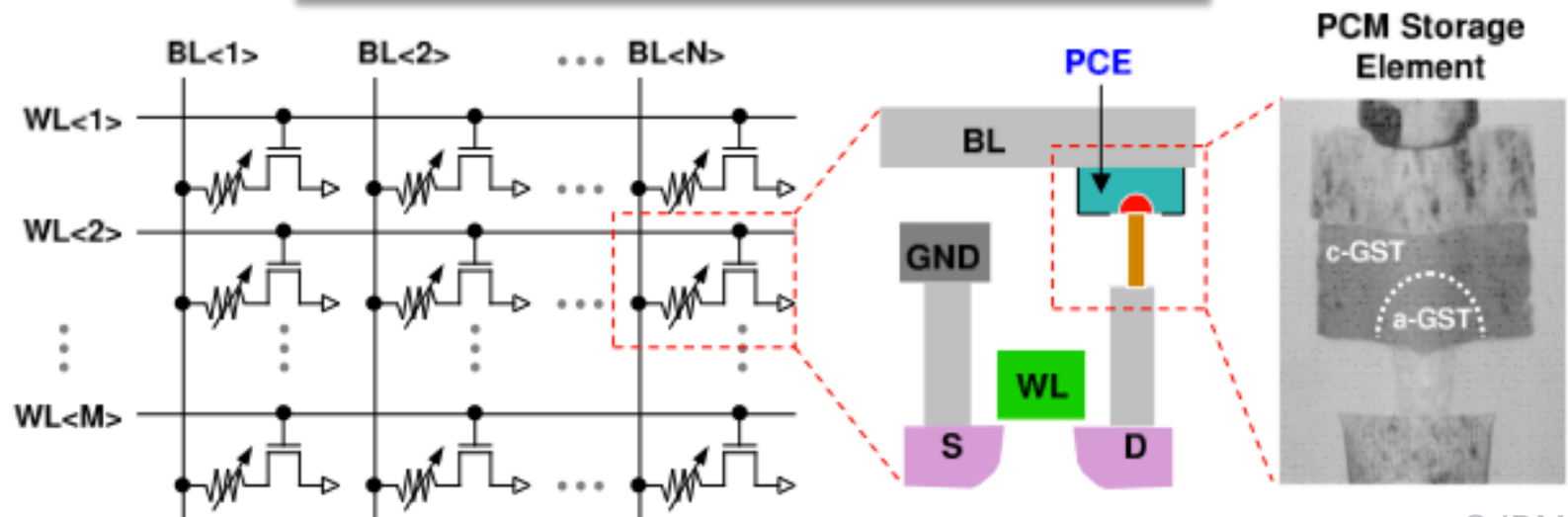
**Neue Ansätze**

# PCM-RAM – Einführung

## PCM = Phase Change Memory

- Aktuell bester Kandidat DRAM und Flash langfristig zu ersetzen
- Veränderung der kristallinen Struktur im Speicher durch Erhitzung/Abkühlung ► Widerstandsänderung des Material

**Widerstand = Information**



© IBM

# PCM-RAM – Überblick

## Vorteile

- Skaliert gut bei  $< 20\text{nm}$
- Nicht flüchtig
- Geringer Energieverbrauch
- Hohe Lesegeschwindigkeit
- Sehr hohes  $Q_{\text{krit}}$  ► Fehler durch Strahlung fast ausgeschlossen

## Nachteile

- Begrenzte Schreibgeschwindigkeit
- Begrenzte Anzahl von Schreibzyklen (momentan  $\sim 10^8$  Zyklen)
  - danach bleibt das letzte Bit „stecken“

# Zwischenstand

1. Fehlerhäufigkeit „in the field“ ist hoch
2. Es dominieren permanente Fehler
3. Fehler im Adressbereich des Betriebssystems häufig
4. Neue Technik wie PCM erzeugt fast ausschließlich permanente Fehler

**Was kann man tun? ...**



# Speicherfehler

Heute + in Zukunft



# Gegenmaßnahmen

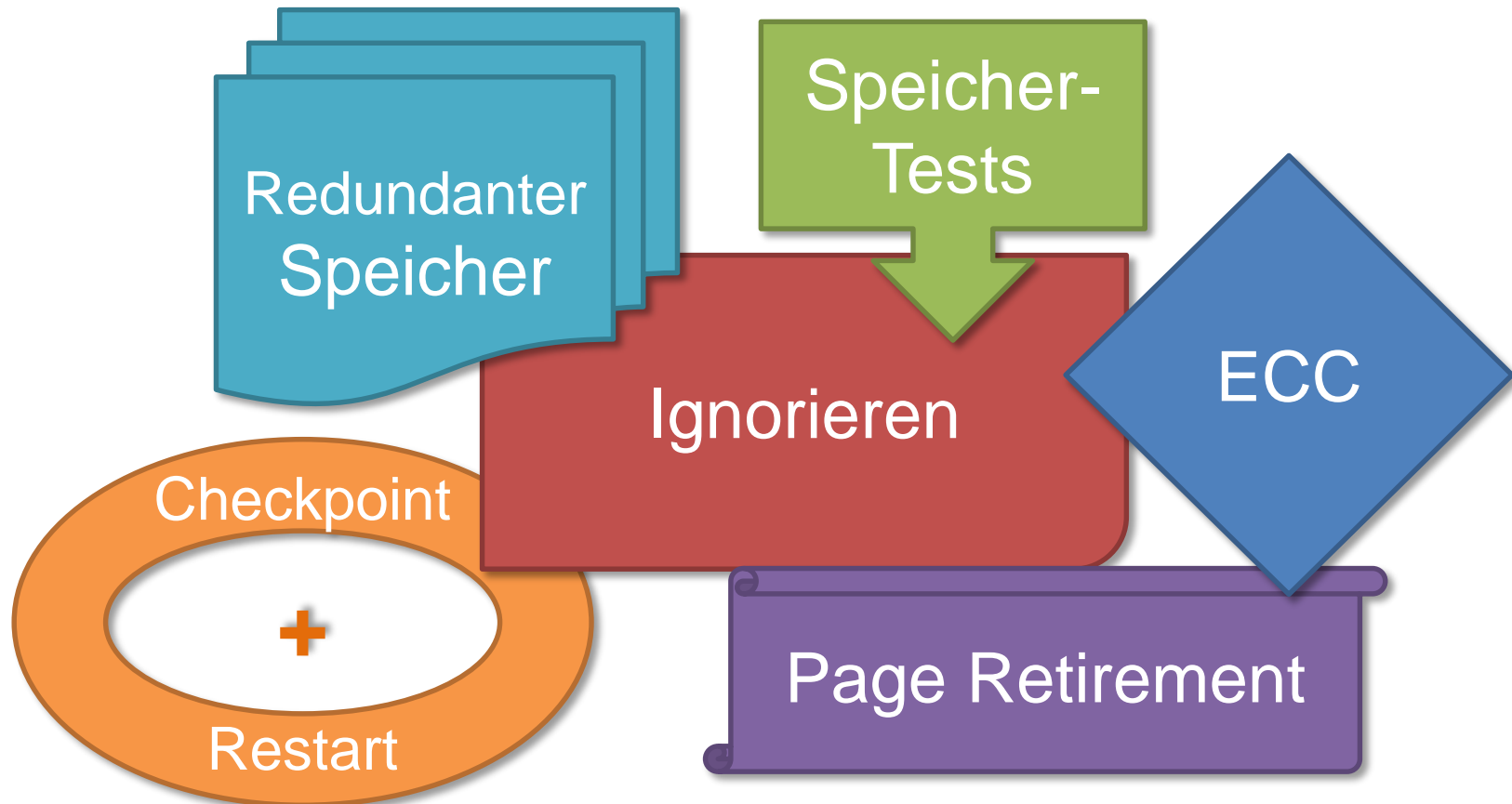
Stand der Technik



# Neue Ansätze

Software + Hardware

# Gegenmaßnahmen – Überblick



**Speicherfehler**

# Gegenmaßnahmen

Stand der Technik

**Betriebssystem**

- Speichertest -

**Hardware**

- ECC -

**Neue Ansätze**

# Stand der Technik – Speichertests

## Speichertests

- RAM wird durch Tool mit Bitmustern beschrieben...
- ... danach ausgelesen
- Bekanntester Vertreter: **Memtest86+**

### Vorteile...

- Test des gesamten Speichers
- Gut darin permanente/sporadische Fehler zu entdecken

Memtest86+ v4.00 | Pass 49% | Test 46% | Test #7 [Random number sequence]

1 Cache: 32K 31865 MB/s | Test #7 [Random number sequence]

2 Cache: 3072K 15194 MB/s | Testing: 180K - 2047M 2046M

3 Cache: None | Pattern: b45cafe2

Memory : 2046M 2326 MB/s |

Chipset : Intel GM965/GL960 (ECC : Disabled) - FSB : 199 MHz - Type : DDR-II

Settings: RAM : 332 MHz (DDR664) / CAS : 5-5-5-15 / Single Channel

WallTime	Cached	RsvdMem	MemMap	Cache	ECC	Test	Pass	Errors	ECC	Errs
0:10:42	2046M	76K	e820	on	off	Std	0	34	0	0

Tst	Pass	Failing Address	Good	Bad	Err-Bits	Count	Chan
2	0	0004a1393e0 - 1185.2MB	00000000	00000400	00000400	1	
4	0	0004a1393f0 - 1185.2MB	45a2d44d	45a2d04d	00000400	2	
6	0	0004a1393e0 - 1185.2MB	01000000	01000400	00000400	23	
6	0	0004a1393f0 - 1185.2MB	00000400	00000000	00000400	24	
6	0	0004a1393e0 - 1185.2MB	00000100	00000500	00000400	26	
7	0	0004a1393f0 - 1185.2MB	76d15652	76d15252	00000400	29	
7	0	0004a1393e0 - 1185.2MB	e32b9b2e	e32b9f2e	00000400	31	
7	0	0004a1393f0 - 1185.2MB	d1e9dc46	d1e9d846	00000400	33	

(ESC)Reboot (c)onfiguration (SP)scroll\_lock (CR)scroll\_unlock

CC Wikimedia.org

### ...Aber

- System während Test blockiert
- Nicht alle permanenten/sporadischen Fehler sicher erkennbar
- Transiente Fehler prinzipiell nicht abgedeckt

# Stand der Technik – ECC

## ECC = Error-Correcting Code

- Nutzung von Informationsredundanz
  - ▶ zusätzliche ECC-Bits zur Erkennung und Korrektur von Fehlern
- Häufig: **DEDSEC** = **D**ouble **E**rror **D**etection **S**ingle **E**rror **C**orrection
- Im Wachstum: Erweitertes **ECC** (*Chipkill, enhanced ECC, ...*)

### Vorteile

- Oft einfach / flexibel einsetzbar
- Schutz transparent während des Betriebs
- Wichtiges Tool für erweiterte Maßnahmen



© Samsung

### Nachteile

- Spezielle Hardware nötig
- Zusätzlicher Overhead
- Sehr gut für transiente Fehler, schlechter für permanente Fehler

# Speicherfehler

Heute + in Zukunft



# Gegenmaßnahmen

Stand der Technik



# Neue Ansätze

Software + Hardware

**Speicherfehler**



**Gegenmaßnahmen**



# Neue Ansätze

Software + Hardware

**Betriebssystem**

- RAMpage -
- Kitten -

**Hardware**

- DRM -

# Linux RAMpage

Schirmeier et al. (2013), TU Dortmund

## Ziele

- Kontinuierliche, automatische Speichertests in Linux
  - ▶ Test im Hintergrund
- Automatisches Poisoning / Page Retirement

## Herausforderung

- Speicherabdeckung
  - Test von Speicherteilen, die aktuell von Programmen oder Betriebssystem belegt werden
  - Testen beliebiger *physischer* Speicherbereiche
- ▶ In Linux so nicht vorgesehen



# Linux RAMpage – Übersicht

## 2-Teilige Struktur

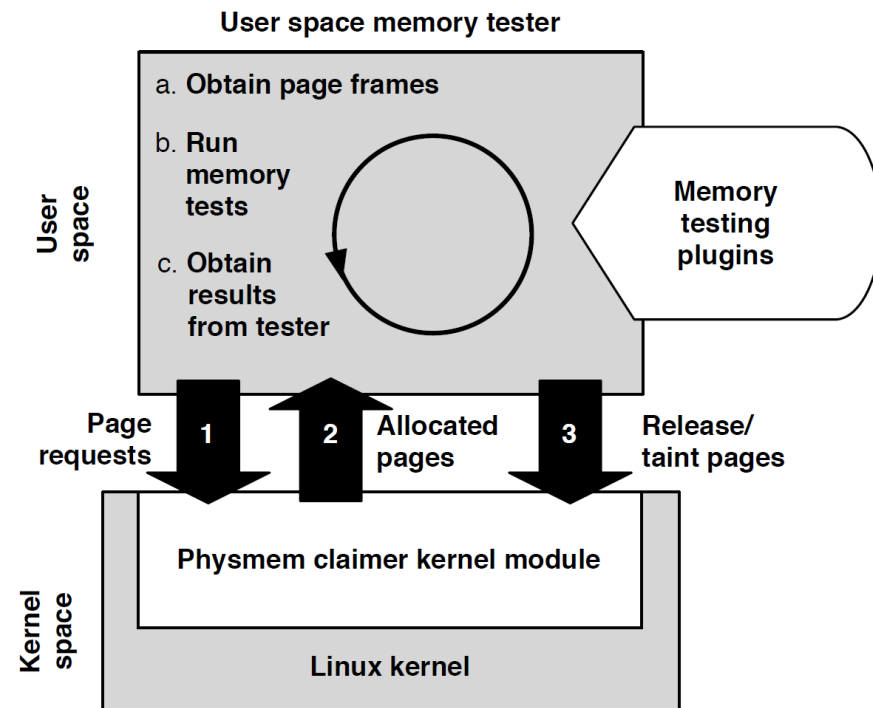
### 1. *User Space Memory Tester*

Hauptmodul, macht Tests

### 2. *Physmem Claimer Module*

„erschleicht“ belegten Speicher

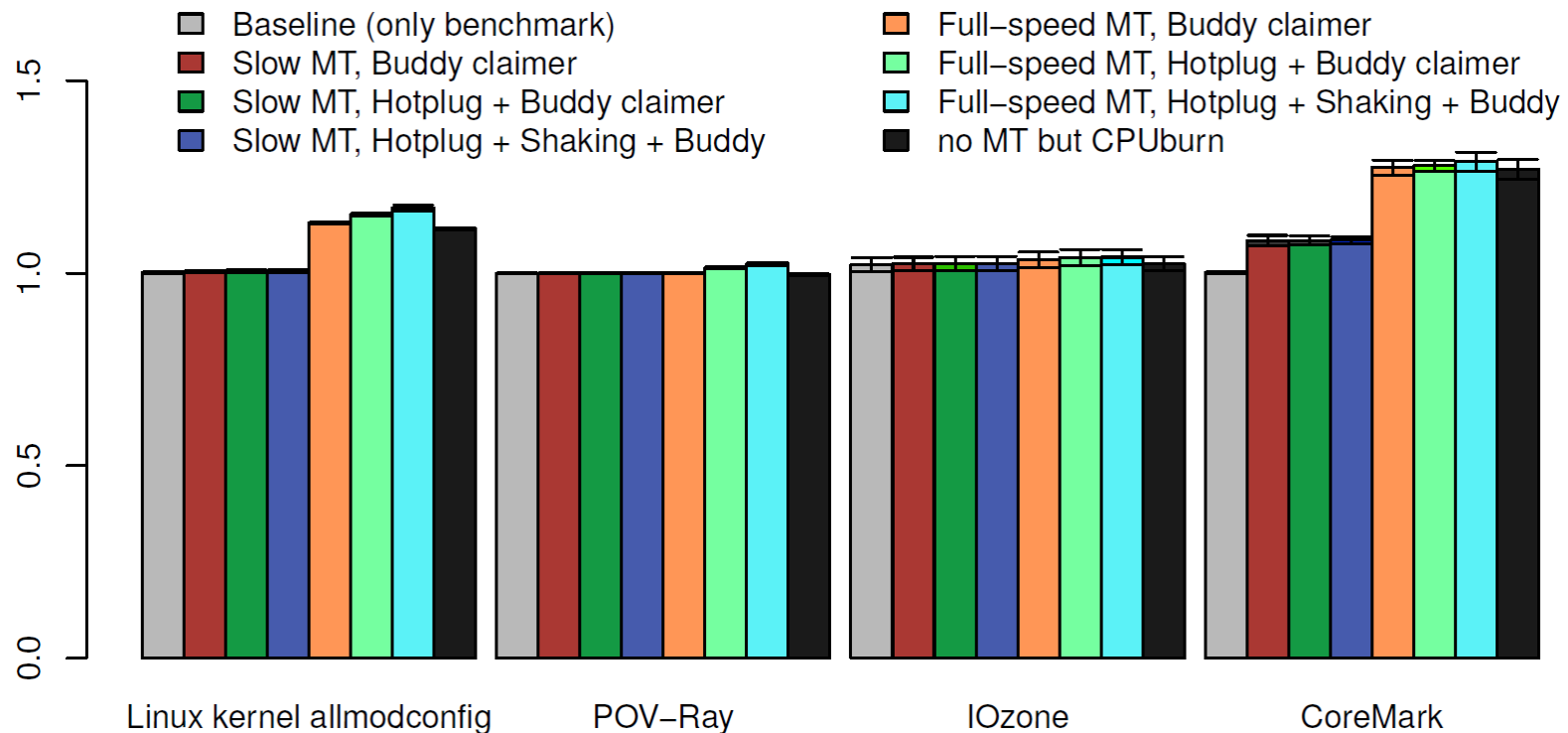
- Linux page allocator (*buddy* allocator)
- memory hotplugging
- hwpoison (*shaking*)



Schirmeier et al. 2013 „Efficient Online Error Assessment and Circumvention for Linux with RAMpage“

# Linux RAMpage – Leistung

## Rechenzeiten bei kontinuierlichem Speichertest



Schirmeier et al. 2013 „Efficient Online Error Assessment and Circumvention for Linux with RAMpage“

# Linux RAMpage – Ergebnisse

## Ergebnisse

- Kontinuierlicher Test (fast) des gesamten Speichers
- Komplettes Testset von Memtest86+
- Kaum Leistungsminderung
- Markierung defekter Speicherstellen
  - ▶ „graceful degradation“
- **Aber:** Kernel-Speicher kann nicht getestet werden
  - ▶ **~95% aller Speicherfehler werden erkannt**

**Speicherfehler**



**Gegenmaßnahmen**



# Neue Ansätze

Software + Hardware

**Betriebssystem**

- RAMpage -
- **Kitten** -

**Hardware**

- DRM -

# Fehlerharte Betriebssysteme

## Zur Erinnerung

- Betriebssystemspeicher ist besonders häufig von Speicherfehlern betroffen
- Fehler hier häufig fatal, da durch Anwendung nicht schützbar

**Kann ein Betriebssystem Speicherfehlern ein besonders schlechtes Ziel bieten?**

# Kitten

## Untersuchung der HPC-Betriebssysteme „Kitten“ und „Cray Linux Environment“

Ferreia et al. (2012)

Sandia National Laboratories

### Hervorzuheben: „Kitten“

- Leichtgewichtiges Betriebssystem
- Linux-Kompatibel
- Open Source
- Durch „Gast-Betriebssysteme“ erweiterbar



© LG-design.com

# Kitten vs. CLE



© LG-design.com



© vicvapor.com

Kitten	CLE (Linux)
30K SLOC	350K SLOC
max. 64MB Kernelspeicher	900 MB+ Kernelspeicher

# Kitten – Ergebnisse

## Die Ziele von Kitten...

- Klein
- Geringe Komplexität
- Determinismus
- Nur die nötigsten Funktionen Teil des „kritischen“ Kernel
- ✓ Kleiner Overhead, geringe Latenz

## ...harmonieren gut mit Fehlertoleranz

- Geringer Speicherbedarf
- Einfache, stark deterministische Datenstrukturen
- ✓ ECC-Schutz der wichtigsten Kerneldaten:  
Speicherverwaltung, Rücksprungadressen... gut umsetzbar



**Speicherfehler**



**Gegenmaßnahmen**



# Neue Ansätze

Software + Hardware

**Betriebssystem**

- RAMpage -
- Kitten -

**Hardware**

- DRM -

# DRM

## DRM – Dynamisch Replizierter Speicher

Univ. Rochester / Microsoft Research

Ipek et al. (2010)

### Ziele

- PCM-RAM realistisch nutzbar machen
- Wenig Overhead
- Wenig Speicherverlust
- Transparenz
- DDR-Standard kompatibel

# DRM

## Motivation

1. Begrenzte Schreibfestigkeit der Speicherzellen
2. Starke Streuung der individuellen Schreibfestigkeiten
3. Etablierte Systeme für SSDs ungeeignet für RAM

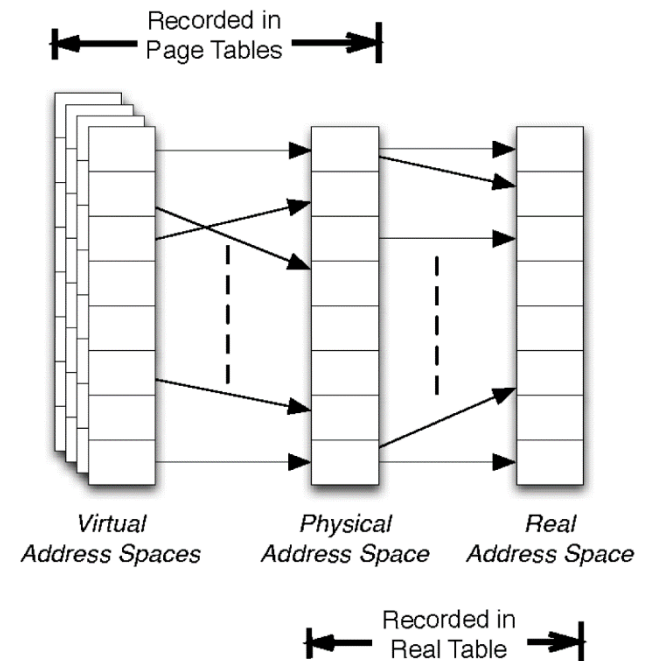
## Bisherige Ansätze

- **Wear-Leveling** ► Grundvoraussetzung, aber unzureichend
- **Multibit ECC** ► zu aufwändig, zu langsam
- **ECC + Page Retirement** ► zu grob

# DRM – Ansatz

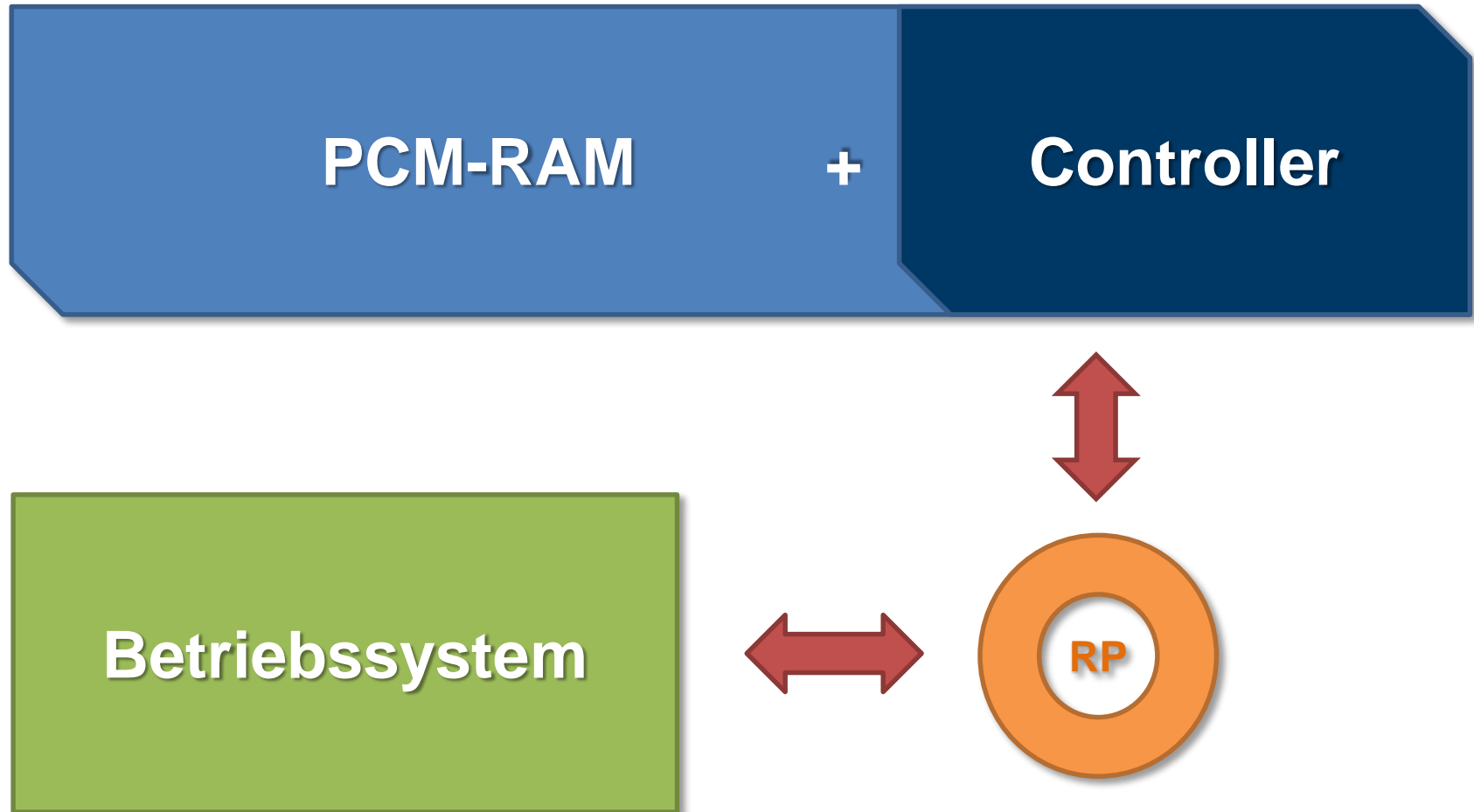
## Lösungsansatz

- ✓ **Dynamische Replizierung von Speicherseiten im Fehlerfall**
  - Gruppierung von beschädigten aber kompatiblen Seiten
  - Zuordnung transparent für das Betriebssystem
  - Sanfter Teilausfall der Speicherleistung / Kapazität („graceful degradation“)

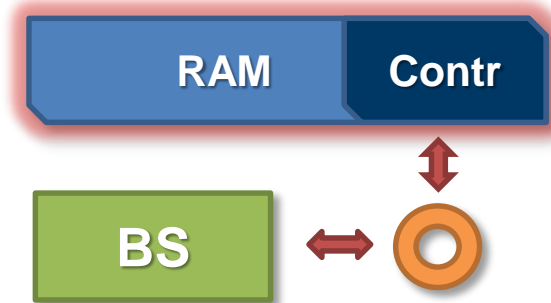
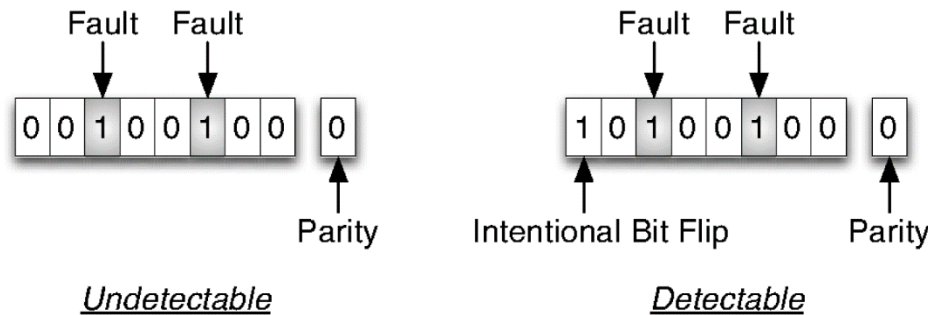


Ipek et al. ASPLOS 2010,  
„Dynamically Replicated Memory“

# DRM – Überblick



# DRM – Daten schreiben



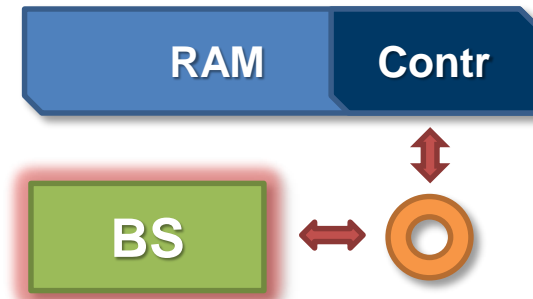
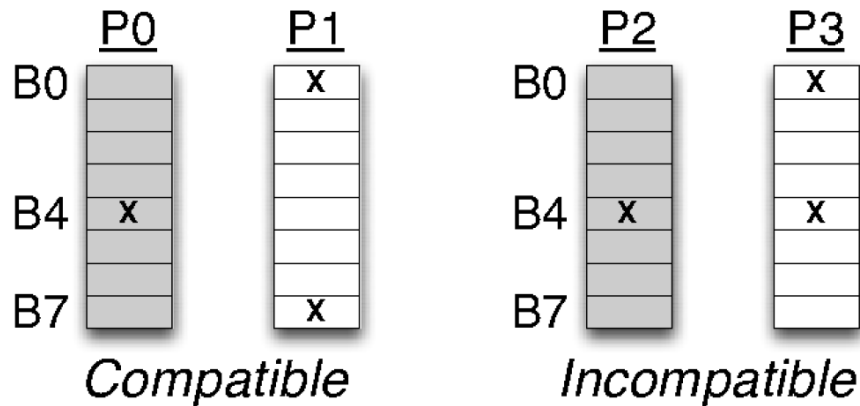
## Speicher

1. Byte wird geschrieben
2. Schreibvorgang wird überprüft ► Fehler
3. Markierung des Beschädigten Blocks (Parity)
4. Meldung des Blocks an Betriebssystem über Ringpuffer
5. Migration der Daten auf neues Seitenpaar aus Ringpuffer

# DRM – Daten schreiben

## Betriebssystem

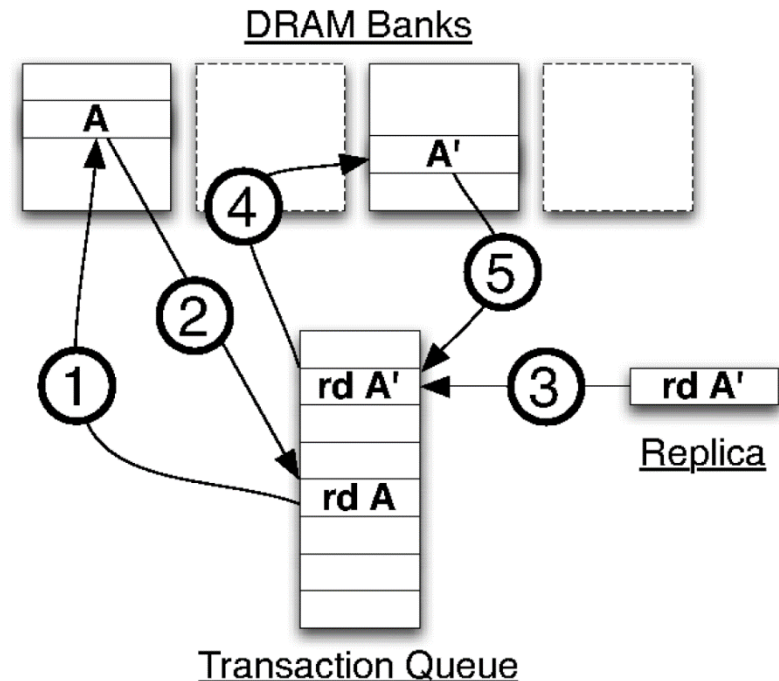
1. Gruppierung beschädigter Seiten
2. Bereitstellung von Replikat-Paaren über Ringpuffer



Ipek et al. ASPLOS 2010,  
„Dynamically Replicated Memory“

# DRM – Daten lesen

1. Anfrage der Daten beim Hauptreplikat
2. Überprüfung ob Block markiert  
▶ ja
3. Planung neuer Anfrage an Replikat
4. Anfrage wird ausgeführt
5. Ausgabe der Daten

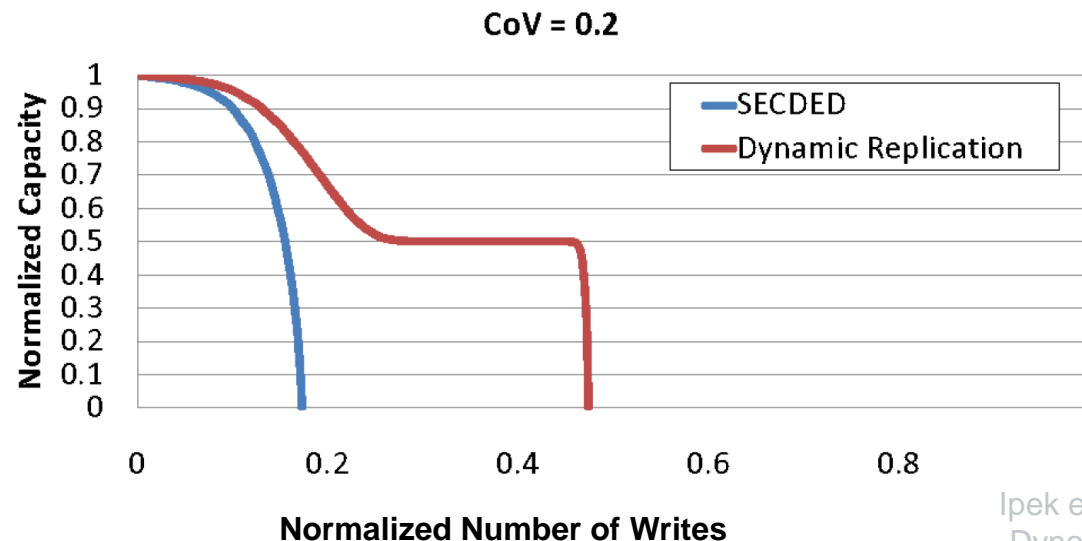




# DRM – Ergebnisse

## Ergebnisse

- Deutlich besseres Ausfallverhalten gegenüber Page Retirement mit SECDED-ECC
- Speicherseiten mit bis zu 160 Bitfehlern noch nutzbar
- Speicher bis zu 40x länger nutzbar
- Leistungssteigerungen durch integriertes Load-Balancing möglich

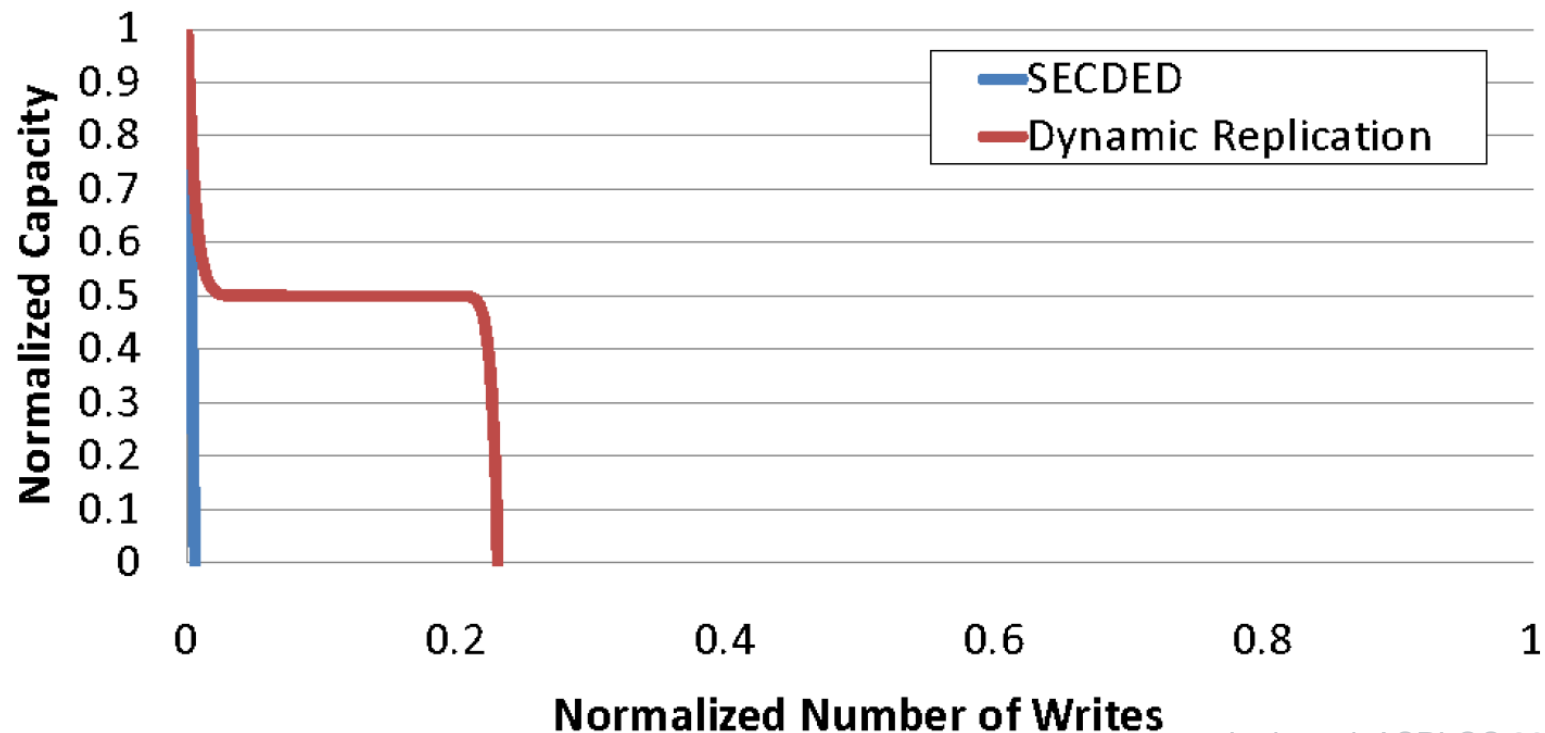


Ipek et al. ASPLOS 2010,  
„Dynamically Replicated Memory“

# DRM – Ergebnisse

## Starke Streuung der Schreibfestigkeit

CoV = 0.3



Ipek et al. ASPLOS 2010,  
„Dynamically Replicated Memory“

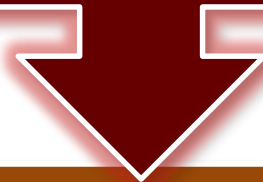
# Zusammenfassung

## Speicherfehler

häufiger als gedacht

Oft permanent

PCM = viele permanente Fehler



## Stand der Technik

ECC + Speichertest



## Neue Ansätze

*RAMpage* – online Speichertest

*Kitten* – klein = weniger anfällig

*DRM* – macht PCM-RAM haltbar

# Zusammenfassung

## Speicherfehler

häufiger als gedacht

Oft permanent

PCM = viele permanente Fehler



## Stand der Technik

ECC + Speichertest



## Neue Ansätze

*RAMpage* – online Speichertest

*Kitten* – klein = weniger anfällig

*DRM* – macht PCM-RAM haltbar

# Zusammenfassung

## Speicherfehler

häufiger als gedacht

Oft permanent

PCM = viele permanente Fehler



## Stand der Technik

ECC + Speichertest



## Neue Ansätze

*RAMpage* – online Speichertest

*Kitten* – klein = weniger anfällig

*DRM* – macht PCM-RAM haltbar

**Vielen Dank für Ihre Aufmerksamkeit!**

