

Ein kostenbegrenzter Ansatz zur Reduktion der transienten Fehlerrate*

Ilia Polian*

John P. Hayes**

Damian Nowroth*

Bernd Becker*

*Albert-Ludwigs-Universität

Georges-Köhler-Allee 51

79110 Freiburg i. Br., Germany

{polian|nowroth|becker}@informatik.uni-freiburg.de

**Advanced Computer Architecture Lab

University of Michigan

Ann Arbor, MI 48109-2122, USA

jhayes@eecs.umich.edu

Zusammenfassung

Transiente Fehler machen sich zunehmend in mikro- und nanoelektronischen Schaltungen bemerkbar. Klassische Härtingsverfahren sind oft mit enormen Kosten verbunden und sind in kommerziellen Produkten nicht wirtschaftlich einsetzbar. Wir stellen zwei Methoden zur selektiven Härtung gegen besonders schwere Fehler vor. Eine Methode identifiziert Fehler, die vom System durch seinen Normalbetrieb in beschränkter Zeit ohne weitere Maßnahmen eliminiert werden. Die andere Methode klassifiziert Fehler je nach Wahrnehmbarkeit ihrer Effekte durch einen menschlichen Endnutzer. Als ein anwendungsspezifisches Beispiel dafür wird eine psychovisuelle Metrik für Bildanwendungen vorgestellt.

1 Einführung

Transiente Fehler (Soft Errors), also vorübergehende Störungen des fehlerfreien Betriebs einer integrierten Schaltung, sind eine neue Herausforderung für Mikro- und Nanoelektronik. Es existiert eine Vielzahl von Härtingsmassnahmen der Schaltungen gegen transiente Fehler, welche jedoch mit enormen Kosten verbunden sind und ausserhalb der Luft- und Raumfahrt sowie der Medizintechnik kaum wirtschaftlich einsetzbar sind.

Bei der *selektiven Härtung* (selective hardening) werden nur die *schwersten* Fehler behandelt, etwa solche mit der grössten Auftretenswahrscheinlichkeit [1, 2]. Die punktuelle Härtung kann unterschiedlich umgesetzt werden; es wurde etwa von Philips vorgeschlagen, die zu härtenden Gatter zu duplizieren [3]. In dieser Arbeit schlagen wir alternativ vor, die Schwere eines Fehlers durch die von ihm verursachte Beeinträchtigung der Systemfunktionalität zu definieren. Die Systemeffekte eines transienten Fehlers werden wie folgt klassifiziert:

1. Kein Effekt
2. Korrekte Daten mit Verspätung berechnet
3. Zeitbeschränkter Effekt (das System kehrt nach einigen Zyklen in den Normalbetrieb zurück)
4. Nicht wahrnehmbarer Effekt (berechnete Daten entsprechen nicht den Referenzwerten, jedoch ist die Abweichung so gering, dass etwa ein menschlicher Endbenutzer sie nicht bemerkt)
5. Permanenter Effekt auf Daten, System läuft weiter
6. Systemcrash

Im nächsten Abschnitt wird auf die Definition der Schwere eines Fehlers bezüglich der Zeitbeschränktheit (Typ 3 der Klassifikation) eingegangen und Ergebnisse

*Diese Arbeit wurde teilweise von der Alexander-von-Humboldt-Stiftung und teilweise von der DFG im Rahmen des Projekts RealTest (BE 1176/15-1) unterstützt.

der entsprechenden selektiven Härtingsstrategie vorgestellt. Im Abschnitt 3 wird für bildverarbeitende Systeme, z.B. Video, eine psychovisuelle Metrik auf der Basis des JPEG-Verfahrens vorgestellt, die sich zur Analyse der Wahrnehmbarkeit des Fehlereffekts (Typ 4) eignet. Als Fehlermodell wird das Flip-to-0/1-Modell angenommen, bei dem eine Leitung der Schaltung für genau einen Taktzyklus ihren logischen Wert auf 0 (1) ändert. Zur Vereinfachung nehmen wir an, dass alle Fehler gleich wahrscheinlich sind.

2 Zeitbeschränkte Effekte

Definition: Ein Schaltkreis C ist *zeitbeschränkt fehler-tolerant* mit Toleranzperiode k und Selbstheilungswahrscheinlichkeit p_{sh} für eine flip-to-0/1-Fehlermenge F , wenn der Zustand des fehlerfreien Schaltkreises und des mit einem beliebigen Fehler aus F behafteten Schaltkreises unter Annahme gleichverteilter Eingaben nach k Schritten mit Wahrscheinlichkeit p_{sh} oder höher übereinstimmen. Wir schreiben kurz, der Schaltkreis sei (F, k, p_{sh}) -ZFT und (k, p_{sh}) -ZFT, wenn die Fehlermenge aus dem Zusammenhang klar ist. \square

In [4] wird nachgewiesen, dass ein Motion Estimator eines MPEG-Systems (96, 1.0)-ZFT für über 70% seiner Fehler ist. Für diese Fehler wird er also unabhängig von der Eingabefolge und dem Startzustand in den fehlerfreien Zustand zurückkehren. Die Schwere eines Fehlers, für welchen der Schaltkreis (k, p_{sh}) ist, hängt von der Anwendung ab. Generell kann man sagen, dass Fehler mit niedrigem k und hohem p_{sh} weniger kritisch sind als andere Fehler.

Die Aufgabe der selektiven Härtung besteht darin, eine gegebene Schaltung (k, p_{sh}) -ZFT zu machen. Wir nehmen an, dass durch die selektive Härtung gegen einen Fehler dieser Fehler vollständig eliminiert wird. Es werden solange Fehler ausgewählt und gehärtet bis die Wahrscheinlichkeit $p_{err}(k)$, dass ein Fehler nach k Taktzyklen immer noch im Systemzustand ist, den Wert $(1 - p_{sh})$ unterschreitet. Die Kosten der selektiven Härtung (Anzahl der gehärteten Fehler und ihr prozentualer Anteil in der

k	$p_{ziel} = 0.9$		$p_{ziel} = 0.99$		$p_{ziel} = 0.999$		$p_{ziel} = 0.9999$	
	Kosten	%	Kosten	%	Kosten	%	Kosten	%
1	50	8.39	351	58.89	508	85.23	553	92.79
2	0	0	186	31.21	403	67.62	492	82.55
3	0	0	99	16.61	293	49.16	457	76.68
4	0	0	23	3.86	188	31.54	368	61.74
5	0	0	0	0	146	24.5	252	42.28
6	0	0	0	0	102	17.11	208	34.90
7	0	0	0	0	42	7.05	182	30.54
8	0	0	0	0	0	0	155	26.01
9	0	0	0	0	0	0	116	19.46
10	0	0	0	0	0	0	67	11.24
11	0	0	0	0	0	0	5	0.84

Tabelle 1: Selektive Härtung für Schaltkreis s298

Fehlermenge) sind in Tabelle 1 exemplarisch für Schaltkreis s298 zu sehen. Die Ergebnisse zeigen, dass die selektive Härtung in der Tat eine kostengünstige Methode ist, eine gegebene Fehlerresistenz zu erreichen, wenn Verhaltensabweichungen innerhalb einiger weniger Taktzyklen akzeptabel sind.

3 Nicht wahrnehmbare Effekte

Ein Fehlereffekt ist nicht wahrnehmbar, wenn der Schaltkreis zwar Werte berechnet, die mit den Referenzwerten nicht übereinstimmen, der menschliche Endnutzer jedoch den Unterschied nicht bemerken würde. Auf die Härtung gegen solche Fehler kann i.d.R. verzichtet werden, insbesondere wenn sie auch zeitbeschränkt sind.

Das verlustbehaftete JPEG-Bildkompressionsverfahren erreicht eine hohe Datenkompression, indem „unwichtige“, also von einem Menschen schlecht wahrnehmbare Bildinformationen verworfen werden [5]. Unterschiedliche Qualitätsstufen werden unterstützt, wobei mit sinkender Qualität mehr Bilddaten als unwichtig eingestuft werden. In diesem Abschnitt wird die Schwere eines Fehlers in der bilderzeugenden Hardware mit Hilfe des psychovisuellen Modells des JPEG-Verfahrens berechnet.

Aus Platzgründen wird an dieser Stelle das Modell intuitiv beschrieben. Für einen Fehler f und eine feste Eingabe(sequenz) werden die vom fehlerfreien bzw. fehlerbehafteten Schaltkreis erzeugten Bilder B und B^f berechnet. B und B^f werden der JPEG-Kompression mit unterschiedlichen Qualitätsstufen unterzogen. Die Schwere des Fehlers wird als die kleinste Qualitätsstufe definiert, für welche die komprimierten Bilder übereinstimmen.

Zum Beispiel wird bei JPEG-Komprimierung mit Qualitätsstufe 1 davon ausgegangen, dass nur psychovisuell irrelevante Bilddaten eliminiert werden, d.h. ein menschlicher Betrachter würde ein so komprimiertes Bild vom Original nicht unterscheiden. Sollte ein Fehler (für eine feste Eingabe) Schwere 1 haben, so müssen Bilder B und B^f auf allen psychovisuell relevanten Daten übereinstimmen. Die Bilder sind also für einen Betrachter nicht unterscheidbar und der Fehler unkritisch. Je höher die Schwere eines Fehlers ist, desto gravierender sind die Unterschiede auf psychovisuell relevanten Bilddaten. Bei mehreren möglichen Eingaben wird das Maximum gebildet.

Die einfachste Metrik aus der Literatur, das *Threshold testing* [6], ist für Bildanwendungen nur bedingt geeignet, da sie die menschliche Perzeption nicht berücksichtigt. Akzeptable Fehler in bildverarbeitenden Schaltkreisen wurde in [7] betrachtet. Auf dem Konzept der Error Tolerance [8] aufbauend wurden permanente, aber keine transienten Fehler betrachtet.

Leider stehen uns derzeit keine Beispielschaltkreise aus der Bildverarbeitungsdomäne zur Verfügung. Daher haben wir eigene Schaltkreise aus Bildern der USC-SIPI Image Database (<http://sipi.usc.edu/database/>) zur Verwendung als vorläufige Benchmarks erzeugt. Zu einem Bild haben wir einen Schaltkreis entworfen, der weder Eingänge noch Zustände besitzt und an seinen Ausgängen konstant die Pixelinformationen des Bildes bereitstellt.

Wir haben für jeden Schaltkreis und jeden seiner bitflip-Fehler seine Schwere ermittelt. Die prozentualen Anteile der Fehler mit Schwere 1 bis 16 sind in Tabelle 2 eingetragen; die letzte Zeile aggregiert alle Fehler mit Schwere über 16. Die meisten Fehler haben eine relativ geringe Schwere. Der Anteil schwerer Fehler ist bei Bildern mit höherer Aktivität, d.h. vielen abrupten Helligkeitsunterschieden (aerial und plant) überdurchschnitt-

S	aerial	airplane	chart	clock	moon	plant
1	26.86	40.02	28.57	39.21	28.73	27.72
2	19.50	23.13	39.68	22.28	22.73	20.87
3	11.44	13.23	9.21	13.05	14.70	12.52
4	10.48	8.87	7.08	8.47	11.18	10.49
5	7.24	5.94	6.10	6.46	8.15	7.43
6	5.95	4.45	2.21	3.71	5.50	5.73
7	4.56	1.97	2.26	2.24	3.62	4.45
8	3.65	0.91	1.36	1.39	2.26	3.29
9	2.84	0.52	0.92	1.02	1.34	2.44
10	2.25	0.31	0.71	0.74	0.73	1.65
11	1.60	0.23	0.48	0.52	0.42	1.17
12	1.21	0.12	0.42	0.32	0.28	0.80
13	0.81	0.12	0.29	0.21	0.18	0.52
14	0.53	0.08	0.21	0.13	0.09	0.32
15	0.40	0.04	0.14	0.09	0.04	0.23
16	0.25	0.02	0.11	0.07	0.03	0.13
>16	0.41	0.04	0.25	0.10	0.02	0.22

Tabelle 2: Wahrnehmbarkeit der Fehlereffekte in aus Bildern generierten Schaltkreisen

lich. Sind Fehler der Schwere 4 noch akzeptabel, so müssen für airplane ca. 15% der Fehler und für plant ca. 28% der Fehler selektiv gehärtet werden.

4 Schlußbemerkungen

Die Behandlung von transienten Fehlern mit traditionellen Methoden scheitert an ihren hohen Kosten. Wir haben einen Ansatz zur Auswahl von kritischen Fehlern für das selektive Härtung vorgestellt. Die Auswahl basiert auf dem Einfluss der Fehler auf die Systemfunktionalität. Das Verfahren ist mit existierenden auf Fehlerwahrscheinlichkeit basierenden Methoden kombinierbar. Während Identifikation der für die Systemfunktionalität kritischen Blöcke heute z.T. manuell durchgeführt wird, erlaubt unser Ansatz eine automatische und systematische Vorgehensweise. Die Schwere der Fehler wird mit Hilfe von Zeitbeschränktheit und Wahrnehmbarkeit erfasst. Vorläufige Ergebnisse zeigen, dass die Fehlerrate mit überschaubarem Zusatzaufwand signifikant verringert werden kann. Unsere künftigen Arbeiten werden auf die Integration weiterer anwendungsspezifischer Informationen über den Bildverarbeitungsbereich hinaus, z.B. psychoakustische Metriken für Audioanwendungen, konzentrieren. Ferner wollen wir unsere Verfahren auf realistischen Signalverarbeitungssystemen evaluieren.

5 Literatur

- [1] K. Mohanram and N.A. Touba. Cost-effective approach for reducing soft error failure rate in logic circuits. In *Int'l Test Conf.*, pages 893–901, 2003.
- [2] C. Zhao, S. Dey, and X. Bai. Soft-spot analysis: Targeting compound noise effects in nanometer circuits. *IEEE Design & Test of Comp.*, 22(4):362–375, 2005.
- [3] T. Heijmen and A. Nieuwland. Soft-error rate testing of deep-submicron integrated circuits. In *European Test Symp.*, pages 247–252, 2006.
- [4] I. Polian, B. Becker, M. Nakasato, S. Ohtake, and H. Fujiwara. Low-cost hardening of image processing applications against soft errors. In *Int'l Symp. on Defect and Fault Tolerance*, pages 274–279, 2006.
- [5] V. Bhaskaran and K. Konstantinides. *Image and Video Compression Standards: Algorithms and Architectures*. Kluwer Academic Publishers, kluwer international series in engineering and computer science edition, 1995.
- [6] Z. Jiang and S.K. Gupta. An ATPG for threshold testing: Obtaining acceptable yield in future processes. In *Int'l Test Conf.*, pages 824–833, 2002.
- [7] I. Chong, H. Cheong, and A. Ortega. New quality metric for multimedia compression using faulty hardware. In *Int'l Workshop on Video Processing and Quality Metrics for Consumer Electronics*, 2006.
- [8] M. Breuer. Error-tolerance and related test issues. In *Asian Test Symp.*, page xxi, 2004.