# 软错误率变动对检查点机制的影响

# 贾文涛 张春元

(国防科学技术大学计算机学院 长沙 410073)

**摘 要**随着集成电路工艺进入纳米时代,芯片正面临着软错误的威胁.除了软错误在数量上的威胁外,处理器还面临着由于工艺变动性和电压、温度、位置等工作环境变动而导致的错误率变动性的威胁,即系统的错误率不会一直稳定而是随着时间发生变化.检查点是系统容错的主要机制,它的开销和检查点的间隔密切相关,目前检查点间隔的确定大多是基于恒定错误率的.但是在软错误变动的情况下,自适应检查点的方法比固定方法更能够显著降低检查点开销,它通过预测系统的错误率来确保系统的检查点间隔始终与最优状态接近.但是自适应检查点所能获得的性能改善与错误率变动的具体程度相关.因此本文研究软错误率变动的形式和幅度如何影响检查点的开销.该文开展了如下研究:基于温度、电压、位置等因素对软错误影响的原理,建立了一个包含变动幅度、持续时间等参数的错误率变动的模型;基于错误率变动模型,模拟了在理想情况下自适应检查点机制能够获得的性能改善; 提出了一种基于错误历史预测错误率的方法,从而验证了在实际情况下自适应检查点能够达到的效果.实验结果表明,变动的幅度在3倍以上且持续时间在12.5%以上时,该文方法就能获得实际上的性能改善.

关键词 容错;软错误;错误变动性;检查点;自适应 中图法分类号 TP311 **DOI**号 10.11897/SP.J.1016.2016.01790

### Impact of Soft-Error-Rate Variability on Checkpoint Mechanism

JIA Wen-Tao ZHANG Chun-Yuan

(College of Computer, National University of Defense Technology, Changsha 410073)

Abstract Soft errors are increasingly important threats to the reliability of integrated circuits as reduction of feature sizes to nanometer level. Chips manufactured in advanced technologies show variations in SER(Soft-Error-Rate) caused by variations in the process parameters and operating environment. Ongoing reduction of feature sizes and complexity of operating environment, SER variation is increasingly manifesting. The most popular recovery method is checkpoint, and the intervals of checkpoint can obviously influence performance. However, most ways to determine intervals of checkpoint relying on constant SER. Theoretically speaking, self-adoptive checkpoint which analyze occurrence of errors more carefully and dynamically match checkpoint interval to real time SER can improve checkpoint overhead under variable SER. But benefit of SACP is relative with SER variation, so we have to evaluate impact of SER variability on self-adoptive checkpoint. We study impacts of theoretical variable SER on checkpoint overhead; propose a way to predict SER based errors occurred most currently, showing practical benefits of self-adoptive checkpoint. Results show our method can improve performance in the situation of variation above 3X amplitude and sustained time more than 12.5%.

Keywords fault-tolerance; soft-error-rate; error variation; checkpoint; self-adaptive

收稿日期:2015-05-20;在线出版日期:2015-11-17.本课题得到国家自然科学基金(61402504,61033008,61272145,61103080)、国家"八六三" 高技术研究发展计划项目子课题(2012AA012706)、高等学校博士学科点专项科研基金优先发展领域课题(20124307130004)资助.贾文涛, 男,1985年生,博士研究生,中国计算机学会(CCF)学生会员,主要研究方向为计算机系统结构、容错计算.E-mail: wtjia@nudt.edu.cn. 张春元,男,1964年生,博士,教授,主要研究领域为计算系统结构、并行计算、计算机可靠性、嵌入式系统.

1791

# 1 引 言

软错误是由宇宙辐射粒子或封装材料中的阿尔法粒子诱发的数据位翻转.随着芯片工艺尺寸的缩小和晶体管数目的增多,芯片面临的软错误的威胁日益严峻.相关研究表明 16 nm 下全芯片的软错误 率 SER(Soft-Error Rate)是 180 nm 下的 100 倍<sup>[1]</sup>.

随着器件尺寸缩小至纳米量级,一次粒子事件 引发多单元翻转 MCU(Multiple-Cell Upset)的情 况日益增多,在 65 nm 时就能发生 MCU. 很多粒子 加速试验结果显示, MCU 会占到所有翻转的 50% 以上,甚至还出现了单次 50 位的翻转<sup>[2]</sup>. 这将使得 目前只能纠正1位错误的 ECC 技术难以为继.

由于工艺尺寸的缩小,系统不仅面临 SER 数量 上和 MCU 的威胁,还面临着新的考验——软错误 率不再恒定而是呈现出变动性.这主要是由器件工 艺变动(process variability)和系统工作环境改变造 成的.

晶体管的供电电压、阈值电压、沟道长度等底层 器件的工艺变动不仅会导致高层整个系统的性能和 功耗的变动,还会导致 SER 的变动.相关实验表明 SRAM(Static Random Access Memory)在两种不 同数据状态(0/1)下的 SER 是不相同的,在 90 nm 时 SER 差别可达 4 倍<sup>[3]</sup>.除了 SRAM 型电路的 SER,逻辑电路的 SER 对系统的影响也日益增大. 不同于 SRAM 型电路的对称性,逻辑电路天然具有 很强的差异性.其 SER 因数据状态(0/1)、时钟状态 (HIGH/LOW)的不同而不同,差异分别可高达 10 倍 和 6 倍<sup>[4]</sup>.这些由工艺变动引起的 SER 差别,在系 统运行时就表现为 SER 不会恒定不变而是随着时 间呈现变动性.

SER本身就和电压、温度、辐射环境等密切相 关,而在系统运行时这些因素也会发生变化.如很多 系统为了降低功耗采用 DVFS(Dynamic Voltage and Frequency Scaling)技术,这会使芯片的电压发 生变化,从而也使系统的 SER 发生变动.随着信息 化和数字化的发展,计算系统渗入各行各业,如商 业航天器、智能驾驶系统、智能城市设施等,这些 系统的工作环境会明显地发生变化.如汽车从武 汉到成都,随着位置变化,SER 会提高 3.5 倍,而 飞机从海平面上升到 12 km,其 SER 会在 30 min 内 增大 100 倍<sup>[5]</sup>. 综上所述,随着未来工艺尺寸的持续缩小,电路 制造和设计造成的 SER 变动会日益加剧.同时,随 着计算机系统日益广泛的应用,越来越多的系统将 会面临更加复杂多变的运行环境.因此将来系统的 SER 变动会越来越普遍.

检查点机制是目前应用最为广泛的错误恢复技术之一.检查点机制的运行如图 1 所示,在应用执行时系统会周期性地保存执行状态,一旦发生错误就返回之前保存的检查点继续执行.



图 1 检查点机制示意图.应用运行的时间 T<sub>solve</sub> 分为若干 时间片 Tex,每个时间片结束后保存检查点状态, Tov 是保存的时间开销.错误 X 发生后,就可以恢复 最近已保存的检查点状态并继续执行应用,这避免 了整个应用的重新执行

检查点的开销定义为 $(T_{total} - T_{solve})/T_{solve}$ ,它和 检查点间隔 *Tex* 的长短相关.根据 Daly<sup>[6]</sup>的检查 点模型,我们模拟了在不同检查点间隔下检查点的 开销.从图 2 可以看出在 *MTTE*(Mean Time To Errors)为100min时最优检查点间隔为21min,在最 优检查点间隔时开销最小.如果 *MTTE*从 100min 变动到10min或1000min,而检查点间隔不跟着变 化时,那开销就变成550%或12%;而检查点间隔跟 着变化到最优时,开销分别为237%或7%.因此从 理论上讲,基于 *SER* 动态调整检查点间隔的自适应 检查点 SACP(Self-Adaptive CheckPoint)能极大地 降低容错开销.



图 2 检查点性能开销和检查点间隔的关系

ACR<sup>[7]</sup>和 FTCS<sup>[8]</sup>也提出了与 SACP 类似的方法,即根据系统的错误率动态调整检查点间隔. SACP 的思路是:即使错误发生的总次数不变,通过 更详细地分析错误发生的局部特性,动态地调整检 查点间隔以使其和错误率更匹配,也可以降低容错 开销. SACP 的优势与 SER 变动的幅度、变动持续 的时间息息相关.但是,导致变动的因素众多且对变 动的影响也千差万别,因此需要详细评测.

本文的贡献是:

(1)根据温度、电压、位置等因素对 SER 影响的原理,建立了一个 SER 变动的模型,它包括了变动幅度、变动持续时间等参数.

(2) 基于 SER 变动模型,模拟了在理想情况下 自适应检查点机制能够获得的性能改善,同时分析了 SER 变动模型中的每个参数对检查点开销的影响.

(3)提出了一种利用最近的错误历史来预测实时错误率的方法.验证了在实际情况下自适应检查 点能够达到的性能改善.

### 2 背 景

本文的背景包含 SER 变动、检查点机制和自适 应检查点. SER 可用平均的错误发生时间 MTTE 衡量,也可用单位时间发生的错误数来衡量,本文选 择后者,即 SER=1/MTTE.

#### 2.1 软错误率的变动

SER 是由宇宙辐射粒子或封装材料的杂质放射的阿尔法粒子诱发的数据位翻转,计算方式为<sup>[9]</sup>

 $SER = F \times A \times K \times e^{-Q_{crit}/Q_s}$ ,

其中, 宇宙辐射通量 F 与海拔高度、纬度、太阳周期、防护材料厚度等因素相关. A 是电路对 SER 敏感的面积. Q<sub>crit</sub>是数据位翻转需要的临界电荷, 与电容、电压成正比, 也与温度有关. Qs 和常数 K 是电路相关的参数. 根据目前掌握的相关文献, 对 SER 有影响的外部环境因素如表 1 所示.

维度、太阳周期等因素对 SER 影响较复杂,本文 不予研究.由表1可见,对于变动因素 X∈{Al,TH, V,T},当每个因素单独变化时,SER 近似等于

$$SER(X) = SER_0 \times e^{C_x \times (X - X_0)}$$
(1)

其中,X0和 SER0为已知的参考值; Cx 为常数,表示每单位 X 的变动造成的 SER 提高或减低的倍数.例如,高度每升高 1 km, SER 增大  $e^{0.8}$ 倍,所以 Cx=0.8.Cx<0表示 SER 随 X 的增大而减小,如 电压、防护材料厚度等.

表 1 影响 SER 变动的因素

可变因素	影响方式	变动幅度 SER_MAX/SER_MIN
高度(AL)	$SER0 * e^{0.9x - 0.03x \times x[5]}$	超过 100(0~12 km)
维度(LA)		3(北纬 20~60 N)
混凝土厚度(TH)	SER0 *e <sup>-2.7x[10]</sup>	$15(0 \sim 1  \text{m})$
太阳活动		1.3 <sup>[10]</sup> 每天之中 1.6 <sup>[10]</sup> 每月之中
电压(V)	$SER0 * e^{-1.3x+0.3[11]}$	3(0.4~1.0V)@45nm
温度(T)	≈指数形式[12]	$32*(21\sim\!125^\circ\!\mathbb{C})$

注:\*温度对错误率的影响除了通过Q<sub>erit</sub>外,还有噪音等多种途径,这 里采用总体上的数据,在21℃基础上每增加10℃错误率加倍<sup>[13]</sup>.

综上所述,由于环境因素导致的 SER 变动幅度 有几倍至百倍,再加上制造工艺和设计变动导致的 SER 差异,SER 的变动幅度会达到千倍.

#### 2.2 检查点机制

检查点机制中的一个重要问题是程序执行多久 才保存状态,这称为检查点间隔的确定.Daly<sup>[6]</sup>根据 图 1 所示的检查点运行方式,将总的运行时间分为 应用运行时间、检查点保存时间、错误恢复时间和重 复计算时间,并对每个部分分别建立数学模型.在总 的运行时间最小的前提下,Daly 求解数学模型得到 最优检查点间隔的计算方法.

$$Tex = \begin{cases} \sqrt{\frac{2\delta}{SER}} \left[ 1 + \frac{1}{3} \left( SER \times \frac{\delta}{2} \right)^{1/2} + \\ \frac{1}{9} (SER \times \delta/2) \right] - \delta, \\ 1/SER, & \delta \ge 2/SER \end{cases}$$
(2)

其中,δ为检查点保存开销,等于图1中的Tov.式(2) 是分段函数,原因是 Daly 建立的数学模型没有精确 解,只能通过高阶多项式逼近精确解,式(2)上半部 分给出的是基于三阶多项式的近似解.精确解的最 大值是 1/SER,只有在δ大于 2/SER 时精确解才 接近最大值,因此此时直接使用最大值比基于多项 式的近似解更为精确.

Daly 的检查点模型和检查点间隔计算方法是 对 Young<sup>[14]</sup>方法的进一步完善,将 Young 方法的 开销误差从 5%降低为 0.2%.本文需要精确地获得 各种方法在不同 SER 下的开销,因此本文采用了 Daly 的检查点模型和检查点间隔计算方法.

#### 2.3 自适应检查点机制

在高性能计算、云计算等领域系统呈现出日益 增大的复杂性和差异性,如任务的特征、资源的利用 率、系统的错误率等因素都会动态地变化.因此人们 提出了多种自适应的检查点机制,以根据系统的状 态动态地调整检查点间隔,如表 2 所示.

(6)

表 2 自适应检查点间隔的相关研究的比较

方法	自适应的原因	自适应的具体方法
ACR <sup>[7]</sup>	在高性能计算领域中系统的错误率比较符合 Weibull 分布,随着运行时间的增大 SER 会下降.	用参数估计的方法获 得Weibull分布的具体 参数,从而计算检查点 间隔.
FTCS <sup>[8]</sup>	网格计算中的资源各异, 不同资源的错误率不同.	由每个资源的失效率 确定检查点间隔,失效 率=已发生的错误数/ 已运行时间.
MFCA <sup>[15]</sup>	网格计算中资源的利用率 和错误率都动态变化.	当 MTTE 大于任务剩 余执行时间时,检查点 间隔增加一个单位量.
本文	器件工艺变动和操作环境 的改变会导致系统的 SER 发生复杂变动.	根据系统的实时错误 率确定检查点间隔.利 用错误的历史来预测 实时错误率.

本文与之前工作的主要不同是本文研究的错误 变动的形式更为复杂、变动幅度更为巨大.

## 3 参数化 SER 的变动

本文的目的是以恒定的 SER 作为参照,研究不同的 SER 变动对检查点性能的影响,为此我们先要参数化 SER 变动. SER(t)表示时刻 t 时的系统软错误率,假设系统初始的错误率为 SER0,即 SER(0)= SER0,则恒定条件下的软错误率 SERc(t)为

 $SER_c(t) = SER(0) = SER0.$ 根据式(1),变动条件下的软错误率  $SER_v(t)$ 为

 $SER_v(t) = SER_0 \times e^{Cx \times (X(t) - X_0)}$  (3) 其中  $X \in \{Al, TH, V, T\}$ ,下面我们来确定 X 随时 间 t 的变动. X 随时间的变动与系统的用途、工作环 境等因素密切相关. 通常 X(t)变动可分为跳跃变动 (如电压)和连续变动(如高度、温度). 跳跃变动时我 们用式(4)描述 X(t)相对参考点 X0 的变动:

$$X(t) = \begin{cases} X0, & t\% T \le T - Tv \\ X0 + a, & t\% T > T - Tv \end{cases}$$
(4)

其中,参数 a 为X(t)在参考点 X0 基础上的具体变 动幅度;a>0 表示在 X0 基础上增大,a<0 表示减 小;T 是X(t)变动的周期,式(4)通过取模运算实现 了X(t)以T 为周期进行变动;Tv 表示变动幅度 a持续的时间,它的范围为  $0 \le Tv \le T$ ,Tv=0 时相当 于X(t)不变动.图 3 是 a=-0.5 时X(t)跳变的示 例,其中周期T=800,跳变持续时间Tv=400.图 3 中没有给出时间单位,因为各个因素的具体变动的 时间有很大差别,图中只是变动形式的示例.

当连续变动时我们用式(5)表示 X(t)

V(t)

$$\begin{cases} X_{0}, & t_{0}^{0}T \leq T - Tv \\ X_{0} + a \times \sin(\pi(t - T + Tv)/Tv), t_{0}^{0}T > T - Tv \end{cases} (5)$$



图 3 变动因素 X 随时间的变化示例

图 3 也给出了 a = 1 时的 X(t)连续变动. 这里用 sin 函数表示变动,主要是因为 sin 函数是周期函数且其导数也是变动的,这对错误预测是一个挑战.

结合式(3),连续变动的软错误率为  $SER_v(t) = \begin{cases} SER0, & t\%T \le T - Tv \\ SER0 \times e^{A \times \sin(\pi(t - T + Tv)/Tv)}, & t\%T > T - Tv \end{cases}$ 

我们令  $A = Cx \times a$ ,因为 Cx 是单位 X 的变动 幅度引起的 SER 倍数, a 为 X 的具体变动幅度,所 以参数 A 表示具体的 X 变动所引起的 SER 倍数. 如对于高度 Cx = 0.8 表示每升高 1 km, SER 增大  $e^{0.8}$ 倍,若 a = 5 km,则 A = 4 表示最大变动幅度为  $e^4$ 倍.图 4 是 A = -1,A = 3 的连续变动和 A = 0.5 的 跳变的 SERv(t)示例.式(6)中如果 A = 0则表示 SER 不随时间变动.



每个周期 T 中错误发生的总次数是 SERv(t) 在时间 0 到 T 上的积分.从图 4 可知,当 A=3 时错 误总次数大于 A=0(不变动)和 A=-1 时的错误总 次数.由于容错开销和错误发生次数密切相关,而本 文研究的目的是错误变动对容错代价的影响,所以必 须消除错误次数对开销的影响.为此,我们引入参数 K 对错误次数进行均等化. SER 不变动时每个周期 T中发生的错误总数为  $\int_{0}^{T} SERc(t) dt = SER0 \times T$ , 我们用式(7)进行均等化错误次数:

$$\int_{0}^{1} SERv(t) dt \times K = SER0 \times T$$
(7)

所以, $K = SER0 \times T / \int_{0}^{T} SERv(t) dt$ . K 是常数,其值 由 A 和 Tv/T 决定. 错误总数均等化后的 SERv 为 SERv(t) =

 $\begin{cases} SER0 \times K, & t\% T \leq T - Tv \\ SER0 \times K e^{A \times \sin(\pi(t - T + Tv)/Tv)}, t\% T > T - Tv \end{cases}$ (8) 式(8)中 SERv(t)的示例见图 5.



图 5 均等化后 SER 随时间的变动形式示例

本文研究的变动错误率对容错开销的影响都是 以式(8)为基础的,我们再总结一下相关参数.

(1) 平均错误率 SER0. SER0 是参考点的错误率,也是 SERv(t) 均等化后每个周期的平均错误率,它的大小决定着每个周期发生的错误总次数.

(2) 变动幅度 A. 错误率变动幅度的定义为
 SER\_MAX/SER\_MIN=e<sup>|A|</sup>,当 A>0 时为增大变
 动,A<0 时为减小变动,A=0 时不变动.</li>

(3) 变动持续时间 *Tv*. SER 以 *T* 为周期变动, *Tv* 表示错误率变动持续的时间,*T*-*Tv* 则是 SER 不变的时间段.

(4) 变动持续时间的比例 *pn*. 在一个变动周期 *T*中,*pn*=*Tv*/*T* 表示变动的时间比例.

根据变动是否连续和 A 是否大于 0,我们将 SERv(t)变动分为表 3 所示的 4 类.

表 3 SER 变动的分类

	A > 0	$A{<}0$
X(t)连续变动	连续向上变动	连续向下变动
X(t)跳跃变动	跳跃向上变动	跳跃向下变动

本文的主要工作就是在这4类变动中研究参数 SER0,A,Tv和 pn 对检查点开销的影响.实际的系 统所面临的 SER 变动要比本文模型中的复杂很多, 但是可以看成是多个简单变动的相互叠加.

### 4 错误率预测

错误率通常用一定时间段中发生的错误数来衡量,在错误率恒定时,选择的时间段越长计算的结果 越精确.但是在错误率变动时,却不是这样,图 6 是用 最近的 200 h 发生的错误数来预测当前的 SER.从 图 6 可以看出,预测的 SER 与理论上的 SER 有明 显的延迟,这主要是由于选择的时间段太长了,因此 错误预测的关键是选择合适的时间段.我们把用来 预测 SER 的时间段称为预测窗口 PW(Prediction Window),图 6 的预测窗口为 200 h.



图 6 SER 预测误差示例. 较粗的线是由式(8)所得的理论 上的 SER,散点是根据粗线模拟的错误发生(每个点 代表一次错误,横坐标是错误发生时间,纵坐标是此 时的瞬时 SER),细线是根据散点预测的 SER. 预测 窗口 PW=200 h

错误率预测的误差由两个因素导致:随机性和 延迟性.错误发生的时刻有很强的随机性,而随机性 会导致错误率计算精度的误差.其解决方法是通过 多次错误来平均,平均的次数越多随机性误差就越 小.由于预测错误率使用的是历史数据,即使随机性 误差很小但是错误率可能在这段时间已经发生了变 化,这称为延迟性误差,其解决办法是尽量用最近的 历史数据,所用的历史数据越接近当前时刻,延迟导 致的误差就越小.遗憾的是这两种误差的解决方法 有很大的冲突,前者要求使用更多的错误次数,意味 着要采用更长时间段的历史数据,也就意味着会增 大延迟性误差.这个矛盾在错误率较小时尤为明显. 因此必须兼顾两者的平衡,本文提出了一种动态确 定预测窗口的方法即 DPW (Dynamic Predication Window)算法来预测 SER. 我们首先定义相关的符 号和函数.

t:当前系统时间; numbErr(pw):在预测窗口 pw 中发生的错误数; SER(pw):以 pw 计算的 SER, 等于 numbErr(pw)/pw; RE( $s_1, s_2$ ):两个软错误率  $s_1 \ \pi \ s_2$  的相对误差 $|s_2 - s_1|/s_1$ .

算法1. 基于动态预测窗口预测 SER.

输入:当前系统时间 t,已发生的错误总数 Nt,错误发生的
 时间序列 ErrTime[Nt](用以计算 numbErr(pw))

输出:当前系统的实时 SER

1. MTTE = t/Nt

2. *pw*=10×*MTTE* //初始化预测窗口

- 3. WHILE numbErr(pw)>50
- 4. pw = pw MTTE; //缩小 pw
- 5. END WHILE
- 6. IF  $pw < 10 \times MTTE$
- 7. RETURN SER(pw); //算法结束

8. END IF

9.  $s_1 = SER(pw)$ ,  $s_2 = SER(pw + MTTE)$ ;

10. WHILE  $RE(s_1, s_2) < 1 \perp numbErr(pw) < 50$ 

11. pw = pw + MTTE; //t t pw 同时更新  $s_1 \ \pi s_2$ 

12. END WHILE

13. RETURN SER(pw);

DPW 的目标是在特定误差限制下寻找合适的 PW 从而均衡随机性误差和延迟性误差.我们给出 的随机性误差限制是最大错误数为 50,即错误次数 大于 50 后我们不再扩大 PW.因为此时随机性导致 的误差已经较小(约 0.02)而需要平衡的是延迟性 误差(算法 3~5 行).所以,当算法在第 3 行发现错 误数大于 50 时就缩小 PW 直到错误数小于 50,然 后计算 SER,算法结束.错误次数小于 50 时算法需 要增大 PW,但是同时需要考虑延迟(算法 6~9 行),我们用 PW 扩大前后的 SER 相对误差来确保 延迟性误差较小.当相对误差较大时,说明 PW 扩 大前后的 SER 变动较大,延迟性误差较重要.

DPW 能实现随机性误差和延迟性误差的平衡, 无论对 SER 连续变动还是跳变,无论向上变动还是 向下变动,都具有总体上较好的适应性.

### 5 验证方法

我们通过模拟 Daly<sup>[6]</sup>的检查点机制来评测检 查点的开销.错误的产生是一个 Poisson 过程,我们 通过随机数来模拟它.我们实现了一个基于时间步 的检查点模拟方法,模拟流程如过程1所示.

过程1. 检查点机制模拟验证方法.

输出:应用本身运行时间 T<sub>solve</sub>,检查点开销为(T<sub>total</sub>-

 $T_{
m solve}$  ) /  $T_{
m solve}$ 

WHILE 模拟时间小于检查点运行总时间 T<sub>total</sub>

模拟时间增加一个时间步 step;

如果错误发生,则开始恢复错误;

否则如果错误恢复完成,则重新执行之前的检查点; 否则如果检查点间隔完成,则

 $T_{\text{solve}} = T_{\text{solve}} + Tex$ , 重新设定检查点间隔.

END WHILE

在具体讨论模拟精度前我们先确定检查点机制的基本参数.根据文献[16],本文选取 SER0 为每  $1/100 \min^{-1}$ , $T_{total}=3200 h$ ,检查点保存的时间开销  $Tov=2\min$ ,错误恢复的时间开销  $Tr=4\min$ .周期 T 和 Tv需要根据 SER0来确定,因为在每个变动周期中足够数量的错误次数才能显示出错误率的变动.本文选择周期 T=800 h,Tv=400 h,则 pn=1/2.本文以上述参数的值为基础,研究变动模型中4个参数的变化对检查点开销的影响.

#### 5.1 确定模拟步长和模拟次数

在过程1中影响模拟精度的有:模拟步长 step 的大小和每种情况下模拟的次数.

步长的大小对模拟的精度影响很大. 对错误率 为  $\lambda$  的 Poisson 过程而言,由于每个 *step* 最多发生 一次错误,因此每个时间步中错误发生次数的期望 的相对误差为( $\lambda \times step - 1 \times (\lambda \times step) e^{-\lambda \times step}$ )/ ( $\lambda \times step$ )=1- $e^{-\lambda \times step}$ . 当步长分别为1min,0.1min 和 0.01min 时( $\lambda = SER0 = 1/100 \min^{-1}$ ),误差分别为 0.99%,0.1%和 0.01%.因此我们选择步长为0.1min 以期望得到 0.1%量级的误差.

我们还要确定每种情况下的模拟次数,多次 模拟进行平均可以降低模拟的随机性误差.根据文 献[17],在99%的置信度和0.1%的预期误差下,当 步长为0.1min时,每种情况下需要模拟的步长数 应为6.64×10<sup>9</sup>,除以每次实验的总步长数3200× 60/0.1,所以每种情况下实验的次数约为3500次.

### 5.2 模拟误差分析和优化

本实验中模拟的精度主要体现在两个方面:一 是模拟的错误发生次数的精度;二是模拟的容错开 销的精度.我们通过比较理论计算的值与模拟的值 来分析模拟误差.基于上文确定的检查点的基本参 数,表4列出了理论计算的错误发生次数和容错开 销.理论上总的错误发生次数为 3200×60/100 = 1920.根据 Daly<sup>[6]</sup>的检查点开销模型,可以计算出 不同变动幅度对应的容错开销.

士计	错误发生的总次数			
刀伝	A = 0	A = 1	A = 3	A = 5
理论计算	1920	1920	1920	1920
模拟 step=0.1 min	1918.73	1917.74	1911.71	1906.56
模拟 step=MIN $\{0, 1, 0, 1 \times SER0/SERv(t)\}$	1918.21	1918.97	1918.10	1918.54
	容错开销和模拟的相对误差/%			
士社	容错	开销和模扎	的相对误	差/%
方法	容错 A=0	开销和模找 A=1	k的相对误差 A=3	$\frac{\cancel{k}}{M=5}$
方法 理论计算	容错 A=0 28.00	开销和模找 <u>A=1</u> 27.76	k的相对误差 A=3 27.06	$\frac{\cancel{2}}{\cancel{2}}$
方法 理论计算	容错 A=0 28.00 28.17	开销和模批 A=1 27.76 27.62	以的相对误 A=3 27.06 26.58	$\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel{\cancel$
方法 理论计算 模拟 <i>ste p</i> = 0.1 min	容错: A=0 28.00 28.17 (0.61)	开销和模批 A = 1 27.76 27.62 (0.50)	k的相对误A=327.0626.58(1.77)	
方法 理论计算 模拟 step=0.1 min 模拟 step=MIN{0.1,	容错: A=0 28.00 28.17 (0.61) 28.17	H销和模拟 A=1 27.76 27.62 (0.50) 27.82	<u>人的相对误</u> <u>A=3</u> 27.06 26.58 (1.77) 27.24	

从表4中可以看出,当 SER 不变动时(A=0) 平均错误数为1918.73,与理论值的相对误差只有 0.07%,这说明了5.1节确定的模拟步长和模拟次 数的有效性.但是随着变动幅度 A 增大到5,错误数 下降为1906.56,而相对误差也增大到0.7%.

这主要是因为在变动幅度较大时,会出现在极短时间中发生很多错误而步长精度不足.所以我们 对模拟步长进行了改进,使用了变动步长 step =  $MIN\{0.1, 0.1 \times SER0/SERv(t)\}$ ,即在错误率较大时使用更小的步长.如表 4 所示,在变动步长下错误次数的相对误差在各个幅度 A 下都小于 0.1%. 对容错开销而言也类似,变动步长可将 A=5 时的 模拟误差从 1.85%降低为 0.63%,从而保证在变动 较大时的容错开销的模拟误差小于 1%. 我们对下面4种方法进行了模拟评估.

(1) FIX. 检查点间隔在执行过程中不变,采用 平均错误率 SER0 确定检查点间隔.

(2) SACP\_OPT. 自适应检查点间隔,在每个检查点都采用理论上的 SER 计算检查点间隔,从而确保每个检查点间隔都是最优的. 这是为了评估 SACP 可达到的最好效果.

(3) SACP\_DPW. 自适应检查点间隔,每个检查点采用 DPW 预测的 SER 计算检查点间隔. 这是评估 SACP 实际上能获得的性能改善.

(4) SACP\_FTCS. 自适应检查点间隔,错误预测的方法基于 FTCS<sup>[8]</sup>. SER=已发生的错误数/已运行的时间,这种方法是动态预测 SER 的最基本的方法. 用来与 DPW 进行比较.

### 6 评估结果

为了描述方便我们约定:将"FIX 开销-SACP 开销"称为绝对性能改善,将"(FIX 开销-SACP 开销)/FIX 开销"称为相对性能改善.如不明确说 明,性能改善默认指相对性能改善.

### 6.1 平均错误率 SER0

图 7(a)中的结果显示, FIX 和 SACP\_OPT 的



图 7 不同 SER0 下检查点的性能开销(柱状图左纵坐标)和相对性能改善(散点线图右纵坐标)(|A|=3,Tv=400h,pn=1/2)

性能开销都随着 SER 的减小而降低. SACP\_OPT 的绝对性能改善从 19%迅速降低为 1.3%,但是相 对性能改善从 17%(SER0=1/12.5 min<sup>-1</sup>)降低到 15%(1/100 min<sup>-1</sup>)后趋于稳定. 这是因为在错误率 很小时,检查点的开销主要是检查点保存的开销,而 不是错误恢复的开销. SER 变动时,SACP 的检查 点保存的次数比 FIX 的要少. 这是因为在错误率高 时 SACP 检查点保存次数多于 FIX,而错误率低时 少于 FIX,但总体而言错误率低的幅度要大于错误 率高的幅度. 总之,在错误率较高时 SACP 可以大 幅度的降低开销;即使错误率很低,SACP 在理论上 也可以获得一定程度(>10%)的相对性能改善. SACP\_DPW 在错误率较高时能显著降低开销 (>10%),但是在错误率小于1/200min<sup>-1</sup>后 SACP\_ DPW 能获得的益处将迅速减小(小于5%).这是 因为在错误率很低时,DPW 的预测精度会急剧减 小,无法获得理论上 SACP 的好处. 与 FTCS 相比, DPW 能获得的性能改善总是更大,特别是在 SER0 较小时 DPW 的优势更为明显.

图 7(b)、(c)、(d)也表现出了类似的结果,总体而 言向上变动时 DPW 的优势比向下变动时更为明显.

### 6.2 变动幅度 A

从图 8 中可以看出,随着变动幅度 | A | 的增加, SACP\_OPT 的性能改善会迅速增大到 15%~30%.



图 8 不同变动幅度 A 下检查点的性能开销(柱状图左纵坐标)和相对性能改善(散点线图右纵坐标) (SER0=1/100 min<sup>-1</sup>, Tv=400 h, pn=1/2)

变动幅度 |A| < 1时, DPW不会获得性能改善. 但是随着 |A|的继续增大 SACP\_DPW 获得的改善会快速增大到  $13\% \sim 23\%$ ,其基本上的走势与OPT相符,特别是对图 8(b)和(d). 与 FTCS相比,DPW 在 A > 1后能获得更好的性能改善,随着 A 的增大优势更为明显.

#### 6.3 变动持续时间 Tv

从图 9 可以看出在 *pn* 不变时,无论 *Tv* 如何变 化,*SACP\_OPT* 的性能改善基本保持不变,对图 9 (b)约为 8.5%,其他约为 15%. SACP\_DPW 可改善的性能与 Tv 密切相关, 在 Tv=25h时性能开销还大于 FIX. 随着 Tv的增 大 DPW 的性能改善逐渐接近 OPT. 从接近的速度 上来看,图 9(b)、(d)要快一点(Tv=200h)而图 9 (a)、(c)慢一点(Tv=800h). 从接近的最终效果上 看,图 9(c)、(d)要好一点(差距为 1.1%)而图 9(a)、 (b)相对差一点(差距为 2%). 与 FTCS 相比,DPW 除了在 Tv>800h时的开销相近外,其他情况均比 FTCS 能够获得更多的性能改善.



图 9 不同变动持续时间 Tv 下检查点的性能开销(柱状图左纵坐标)和相对性能改善(散点线图右纵坐标) (|A|=3, SER0=1/100 min<sup>-1</sup>, pn=1/2)

### 6.4 变动持续的时间比例 pn

从图 10 可以看出在 pn 小于 0.5 时, SACP\_OPT 的性能改善迅速增大.当 pn 继续增大时,图 10(a) 的性能改善却开始下降而图 10(b)还是继续增大, 这个原因比较复杂.我们通过误差比率进行说明, 图 11 显示了 FIX 相对 OPT 的误差比率.

某一个时间点 t 的误差比率 ER(t) 定义为 MAX(SER\_FIX, SER\_OPT(t))/MIN(SER\_ FIX,SER\_OPT(t)),当SER\_FIX>SER\_OPT(t) 时,称此时的误差为过量误差(SE\_OV),否则称为 不足误差.

图 11 显示了总误差比率、过量误差比率和过量 误差占总误差的比例.对图 11(a)和(b),总误差比 率都从 1 增加到 3.8(*pn*=6/8),然后开始下降.但 是图 11(a)中过量误差的比例从 100%下降到 91%,而图 11(b)中却从 0 增加到 94%.相同误差 下,过量误差比不足误差对检查点性能的影响要大, 这也与已有的研究结果相符,所以才会出现图 10 中 *pn* 对图 10(a)、(b)的不同影响.

从图 10 可以看出,DPW 的性能改善和 OPT 的比较接近.在 pn>1/8 后 DPW 的性能改善比 FTCS 的更好.对于跳变,其趋势也类似,限于篇幅就不叙述了.





图 10 不同变动持续比例 *pn* 下检查点的开销(柱状图左 纵坐标)和相对性能(散点线图右纵坐标)(|A|=3, *SER*0=1/100 min<sup>-1</sup>,*Tv*=400 h)



图 11 不同变动比例 pn下的总误差比率、过量误差比率 (柱状图左纵坐标)和过量误差的比例(散点线图右 纵坐标)

### 7 总 结

随着器件工艺尺寸的缩小和系统运行环境的日 益复杂,系统的 SER 会呈现出明显的变动性. SER 变动的幅度从几倍到几百倍.检查点是广为使用的 容错技术,其开销和检查点的间隔密切相关,而检查 点间隔的确定又和错误率密切相关.

自适应检查点机制通过预测系统的错误率从而确保系统的检查点间隔始终与最优状态接近.目前的研究很少关注 SER 变动,本文是首次研究 SER 变动对检查点开销的影响.

(1) 基于系统的温度、电压、位置等因素对软错 误率影响的原理,本文建立了包含参数 SER0,A, Tv,pn 的错误率变动的模型. 根据 SER 变动的特 点将 SER 变动分为 4 种类型:连续向上变动、连续 向下变动、跳跃向上变动和跳跃向下变动.

(2)针对4种变动类型,本文研究了模型中每 个参数对检查点性能的影响程度,模拟了在理想情况下自适应检查点机制能够获得的最大性能改善.

(3)本文提出了基于动态预测窗口算法进行错误率的实时预测的方法,验证了在实际情况下自适应检查点能够达到的性能改善.实验结果表明变动的幅度在3倍以上且持续时间在12.5%以上时,本

文方法就能获得实际上的性能改善.

本文通过大量的模拟实验,获得了 SER 变动的 具体形式(SER0,A,Tv,pn)对检查点开销的影响 程度和影响的趋势.实际的系统所面临的 SER 变动 要比本文模型中的复杂很多,但是可以看成是多个 简单变动的相互叠加.因此,本文的研究结果对具 体系统的容错设计具有参考价值,例如系统是否需 要考虑 SER 变动、是否需要采用自适应检查点机 制等.

参考文献

- Borkar S. Designing reliable systems from unreliable components: The challenges of transistor variability and degradation. IEEE Micro, 2005, 25(16): 10-16
- [2] Nicolaidis M. Soft Errors in Modern Electronic Systems. New York, USA: Springer Science and Business Media, 2011
- [3] Heijmen T. Spread in alpha-particle-induced soft-error rate of 90-nm embedded SRAMs//Proceedings of the 13th IEEE International On-Line Testing Symposium. Heraklion, Crete, Greece, 2007: 131-136
- [4] Heijmen T, Roche P, Gasiot G, et al. A comprehensive study on the soft-error rate of flip-flops from 90-nm production libraries. IEEE Transactions on Device and Materials Reliability, 2007, 7(1): 84-96
- [5] Zielger J F, Puchner H. SER-History, Trends, and Challenges: A Guide for Designing with Memory ICs. San Jose, USA: Cypress Semiconductor Corporation, 2004
- [6] Daly J T. A higher order estimate of the optimum checkpoint interval for restart dumps. Future Generation Computer Systems, 2006, 22(3): 303-312
- [7] Xiang N, Meneses E, Jain N, Kale L V. ACR: Automatic checkpoint/restart for soft and hard error protection// Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. Denver, USA, 2013; 1-12
- [8] Mohammed A. A fault tolerant scheduling system based on checkpointing for computational grids. International Journal of Advanced Science and Technology, 2012, 48(1): 115-124
- [9] Hazucha P, Svensson C. Impact of CMOS technology scaling on the atmospheric neutron soft error rate. IEEE Transactions on Nuclear Science, 2000, 47(6): 2586-2594
- [10] JEDEC standards on measurement and reporting of alpha particle and terrestrial cosmic ray induced soft errors in semiconductor devices. Arlington, VA: JEDEC Solid State Technology Association, JESD89, 2006
- [11] Chandra V, Aitken R. Impact of technology and voltage scaling on the soft error susceptibility in nanoscale CMOS// Proceedings of the IEEE International Symposium on Defect

and Fault Tolerance of VLSI Systems. Boston, USA, 2008: 114-122

- [12] Jagannathan S, Diggins Z, Mahatme N N, et al. Temperature dependence of soft error rate in flip-flop designs//Proceedings of the IEEE International Reliability Physics Symposium. Anaheim, USA, 2012: SE. 2. 1-SE. 2. 6
- [13] Yi Hui-Zhan. Research on Low-Power Techniques: Architecture and Compiler Optimization [Ph. D. dissertation]. Graduate School of National University of Defense Technology, Changsha, 2006(in Chinese)
  (易会战. 低功耗技术研究一体系结构和编译优化[博士学位 论文]. 国防科学与技术大学研究生院,长沙,2006)
- [14] Young J W. A first order approximation to the optimum checkpoint interval. Communications of the ACM, 1974, 17(9): 530-531
- [15] Lidya A, Therasa S, Sumathi G, Antony D S. Dynamic adaptation of checkpoints and rescheduling in grid computing. International Journal of Computer Applications, 2010, 2(3): 95-99
- [16] Jones W M, Daly J T, Debardeleben N. Impact of sub-optimal



JIA Wen-Tao, born in 1985, Ph. D. candidate. His research interests include computer architecture and fault-tolerant computing. checkpoint intervals on application efficiency in computational clusters//Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing. New York, USA, 2010: 276-279

- [17] Shubu Mukherjee. Architecture Design for Soft Errors. San Francisco, CA: Morgan Kaufmann Publishers Inc, 2008
- [18] Gerofi B, Ishikawa Y. Workload adaptive checkpoint scheduling of virtual machine replication//Proceedings of the 2011 IEEE 17th Pacific Rim International Symposium on Dependable Computing. Washington, USA, 2011: 204-213
- [19] Zhang Chengye, Deng Shenglan, Ning Hong. A local checkpoint mechanism for on-board computing//Proceedings of the 2012 IEEE International Conference on Information Science and Technology. Hubei, China, 2012: 520-526
- [20] Xiong Lei, Tan Qingping. Data flow error recovery with check pointing and instruction-level fault tolerance//Proceedings of the 12th International Conference on Parallel and Distributed Computing, Applications and Technologies (2011). Gwangju, Korea, 2011: 79-85

ZHANG Chun-Yuan, born in 1964, Ph. D., professor. His research interests include computer architecture, parallel programming, reliable computer system and embedded systems.

#### Background

Soft errors including are increasingly important threats to the reliability of integrated circuits fabricated in advanced CMOS technologies. Checkpoint which is one of most popular method to recovery errors is facing higher overhead as soft-errors-rate is rising in nano-scaled chips. Many works study on decreasing overhead of checkpoints.

Chips manufactured in advanced technologies show variations in *SER* caused by variations in the process parameters. Moreover, *SER* is highly associated with operating environment (temperature, voltage, radiation pressure and so on), which is not always constant. Ongoing reduction of feature sizes and complexity of operating environment, *SER* variation is increasingly manifesting.

So, our problem is how to decrease checkpoint overhead for variable *SER*. The most difference between previous works is that we are focusing on *SER* variation which has rarely been studied. Based on how SER affected by temperature, voltage, radiation pressure and so on, we describe SER Variation with four parameters: average SER, variation amplitude A, variation duration Tv and variation proportion n, with a exact simulation, we find out how each parameter impacts the overhead of checkpoint. We proposal a way to predict SER based dynamic prediction window, study impacts of variable SER on checkpoint overhead showing optimal and practical benefits of SACP. Results show our method can improve checkpoint overhead in the situation of variation above 3X amplitude and sustained time more than 12.5%.

Supported by the National Nature Science Foundation of China under Grant Nos. 61402504, 61033008, 61272145, and 61103080, the National High Technology Research and Development Program of China under Grant No. 2012AA012706, the Research Fund for the Doctoral Program of Higher Education of China under Grant No. 20124307130004.