

上海市教育委员会高校重点教材建设项目
上海市教育委员会 组编

可信计算系统设计和分析

徐拾义 编著

清华大学出版社

上海市教育委员会高校重点教材建设项目
上海市教育委员会 组编

可信计算系统设计和分析

徐拾义 编著

清华大学出版社

北京

内 容 简 介

当前,随着“普适计算”时代的到来,数字计算系统已经渗透到社会的各个领域。对“可信计算和可信计算系统”的设计理论和实践环节的深入研究和开拓必将成为当今数字系统研究的热点。

本书共有 12 章,可以分成两大部分。第一部分为对可信性各类属性和基本知识的详细介绍和分析(第 1~4 章)。其中包括对软硬件系统的故障、错误和失效的定义及性质的形式化描述,并对软硬件系统中的故障、错误作了分析和比较;对可信性的 6 个属性的定义以及评估和计算方法等进行了讨论。第二部分则是在对可信性的主要属性做深入研究的基础上,阐述了提高系统可信性的基本理论、主要技术和实施方法及其他相关知识(第 5~12 章)。其中详细分析和讨论软硬件系统的开发生命周期各个阶段应采取的各种可信性策略和措施,阐明了避错技术、防错技术、排错技术、可测性设计技术(包括冗余与编码技术)、容错设计技术以及故障安全技术等在可信计算系统中的理论、方法和应用实例。

本书吸收了国内外关于可信性理论和技术方面的大量研究成果,是一本集数字计算机软硬件系统于一体、可信性理论和实践并重的著作,适应作为高等院校计算机科学、电子信息、通信工程以及微电子等相关专业高年级本科生和研究生的教材,也可供与可信计算有关的专业人士学习参考。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

可信计算系统设计和分析/徐拾义编著. —北京:清华大学出版社,2006.7

ISBN 7-302-12613-5

I. 可… II. 徐… III. ①电子数字计算机—系统可靠性—系统设计—高等学校—教材 ②电子数字计算机—系统可靠性—系统分析—高等学校—教材 IV. TP33

中国版本图书馆 CIP 数据核字(2006)第 015401 号

出 版 者:清华大学出版社 地 址:北京清华大学学研大厦

<http://www.tup.com.cn> 邮 编:100084

社 总 机:010-62770175 客户服务:010-62776969

组稿编辑:陈国新

文稿编辑:顾 冰

印 刷 者:清华大学印刷厂

装 订 者:三河市春园印刷有限公司

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:21.25 字数:524 千字

版 次:2006 年 7 月第 1 版 2006 年 7 月第 1 次印刷

书 号:ISBN 7-302-12613-5/TP·8064

印 数:1~3000

定 价:33.00 元

前言

随着信息技术领域研究开发的迅速发展,人们对信息技术产品(主要指数字计算系统)的依赖程度愈来愈大。这些技术和产品不仅应用于航空航天、原子核反应堆、军事科学、国防建设以及国民经济各个高科技领域,而且还直接牵涉到人们的生活质量,甚至关系到人类生命、财产的安全问题。因此,当前人们在应用这些产品的同时,必然会提出更高的要求,即除了传统意义上的要求和标准以外,还提出了更重要的评价体系——系统所提供服务的“可信性”(dependability)标准问题,有些专家又进一步提出系统“诚信”(trustworthy)的概念。事实上,对任何一件产品(一个系统)来说,能否为人类提供一种“可信”的服务乃是衡量其总体质量最重要的标准之一。在过去,由于受各种条件的限制,“可信计算系统”是数字计算系统中的阳春白雪,世界上只有少数发达国家在某些特殊要求下,才有可能去开发和应用它们。例如在航天飞机飞行姿态控制系统、原子核反应堆控制系统、导弹防御系统、铁路运输信号控制系统以及一些重症病人监护系统等需要提供十分可靠和可信服务的领域中才会使用它们。相反,这些控制系统中的任何隐患、故障和错误都将对系统和环境带来严重后果,甚至给人类造成巨大灾难。人类在享受着信息技术、数字计算系统提供的各种前所未有的服务和乐趣的同时,却也已经受了由于信息、系统的不可信带来的损失和挑战,并可能为此付出沉重的代价,而这样的例子已经不胜枚举了。因此,人们渴望着有一天可以从可信数字系统中获得更便捷、更优质的服务,而完全免受不良的后果。

当前,“普适计算”(pervasive computing)时代已经来临,既然数字计算系统已经渗透到社会的各个领域,甚至达到“家喻户晓”的程度,那么“可信计算系统”就不再是神秘系统了,而且等待着的将是对可信计算系统的深入研究和规模性开发。它将迅速深入到人类赖以生存各个领域和环境中去。因此,人们期待着有一天可以从可信数字系统中获得巨大的利益和优质的服务,又可以免受不良的后果。

一个数字计算系统的可信性是指该系统提供确实可信服务的综合能力。之所以称它为一种综合能力,是因为衡量这个能力的标准并不是、也不可能是单一的。事实上,它由多方面综合因素确定的评价体系。该体系主要涉及到以下6个属性,即可靠性(reliability)、可用性(availability)、可测性(testability)、可维护性(maintainability)、安全性(safety)以及保密性(security)等。上述诸方面组成了可信性度量的主要属性。

本书最显著的两大特色是两个综合:一是软硬件系统的综合研究;二是6大属性的综合分析。

软硬件系统的综合研究:对数字系统中的软件系统和硬件系统的可信性问题同时进行研究和讨论,特别是提出了两者在可信性理论和实践中的相似性和差异性。长期以来,由于软件系统和硬件系统之间的物理形态和表现方式等存在着很大的差异,软件专业(行业)和硬件专业(行业)的专家、学者以及相关专业学生的研究、开发工作经常是相互独立,分开进

行的。也就是说,软件系统的专家与硬件系统的专家在可信性领域内的研究工作一般都自成系统,软硬件之间少有合作。但是,从数字系统内在的整体观点来看,软硬件系统的总体设计、开发以及系统的可信性理论等是不可分割的统一体,它们在许多方面都是相通的,也是可以互补的。例如,对一个数字系统来说,仅仅讨论硬件系统的可靠性,而忽略了软件可靠性的重要环节,则对整个系统来说,只能认为解决了50%的可靠性问题。另外,即使对硬件工程师来说,也需要了解软件系统的可信性的相关原理。同样,软件工程师更应该懂得硬件系统的可信性基本原理,这样,“软硬结合”就更有利于提高整个系统的可信性。为此,本书有意识地将软硬件系统的可信性问题同时进行研究,通过对两者进行深入的比较,指出了两者的内在联系,也说明它们的差异所在。并且,在某些关键理论和技术方面,还特意同时给出了软件和硬件相关的例子,以同时满足软件和硬件专业工程师们的兴趣和要求。因此,本书无论对软件专业学生(工程师)还是硬件专业学生(工程师)都具有一定的现实意义。

6个属性的综合分析:将上述6个属性作为组成系统可信性度量的基本属性进行综合性的分析讨论。这样对系统的可信性可以有一个全面、客观的评价,同时,也有利于读者对每个属性重要性的理解。其中,可靠性是对系统在一定的环境中能正常完成所规定任务的能力的评价,它的评价标准主要在于系统能保持正常运行的时间长度。可测性则是对一个系统具有的避免故障,防止故障以及检测、诊断和修复故障能力的评价。可用性和可维护性是对系统在具有一定的维护措施之后,随时可提供正常服务的能力的评价。安全性则是在系统一旦发生失效后可能产生后果严重性的评估。而保密性是关于对系统保持其内部所存信息的认证性和完整性能力的评价。本书将对可信性的上述各类属性分别做详细的分析研究,但是,由于保密性属性的评价体系牵涉到信息保存、信息处理、社会安全、人为因素以及其他方面的知识领域,因此不属于本书的讨论范围。

本书共有12章,可以分成两大部分。第一部分为对可信性各类属性和基本知识的详细介绍和分析(第1~4章)。其中包括对软硬件系统的故障、错误和失效等重要概念的基本定义及性质的形式化描述,并对软硬件系统中的故障、错误做了分析和比较。对可信性的6个属性的定义以及评估和计算方法等进行了讨论。第二部分则是在对可信性的主要属性做深入研究的基础上,阐述了提高系统可信性的基本理论、主要技术和实施方法以及其他相关知识(第5~12章)。其中详细分析和讨论了软硬件系统开发生命周期各个阶段应采取的各种可信性策略和措施,阐明了避错技术、防错技术、排错技术、可测性设计技术(包括冗余与编码技术)、容错设计技术以及故障安全技术等在可信计算系统中的理论、方法和应用实例。

本书吸收了国内外有关可信性理论和技术方面的大量研究成果,是一本集数字计算机软硬件系统于一体、可信性理论和实践并重的著作,适合作为高等院校计算机科学、电子信息、通信工程专业以及微电子等专业的高年级本科生和研究生的教材,也可供与可信计算有关的专业人士学习参考。

作者感谢彭成寒、肖全亮、朱伟和徐睿虹在本书撰写和编辑过程中的辛勤劳动和所做的贡献。

徐拾义

2006年5月于上海

syxu@fudan.ac.cn

目 录

第 1 章 数字计算系统可信性及属性	1
1.1 可信性定义及问题的提出	1
1.2 影响可信性计算的主要因素	2
1.2.1 失效的定义及后果.....	3
1.2.2 故障的定义及性质.....	4
1.2.3 错误的定义及传递性.....	6
1.3 可信性的评估测度和提高可信性的措施	9
1.4 小结.....	10
1.5 思考题.....	10
参考文献	10
第 2 章 故障及故障模型	11
2.1 故障模型及分类.....	11
2.2 硬件结构类/功能类故障模型	13
2.2.1 逻辑电路的结构类/功能类故障模型.....	14
2.2.2 CMOS 门电路的结构类/功能类故障模型	14
2.3 软件结构类/功能类故障模型	15
2.3.1 软件结构类故障模型	15
2.3.2 软件功能类故障模型	16
2.3.3 软件功能类故障的一个实例	18
2.4 故障模型的建立标准.....	19
2.4.1 建立故障模型的标准	20
2.4.2 故障模型的不足	21
2.5 小结.....	22
2.6 思考题.....	22
参考文献	23
第 3 章 可信性测度和评估	24
3.1 数字系统的可信性测度.....	24
3.1.1 可信性的数量测度	24
3.1.2 可信性的质量测度	25
3.2 硬件系统的可靠性及其测度.....	25

3.2.1 可靠性函数和失效函数	26
3.2.2 MTTF, MTTR 和 MTBF 的定义	30
3.3 组合系统的可靠性	33
3.3.1 串行系统的可靠性	33
3.3.2 并行系统的可靠性	34
3.3.3 串并/并串系统的可靠性	35
3.3.4 非串行/非并行系统的可靠性	37
3.4 软件系统可靠性及其测度	39
3.5 可测性及其测度	39
3.6 可维护性及其测度	40
3.6.1 维护及其定义	40
3.6.2 可维护性及其定义	41
3.7 可用性及其测度	43
3.8 安全性及其测度	45
3.9 保密性及其测度	46
3.10 可信性的综合评价标准	47
3.11 小结	48
3.12 思考题	49
参考文献	50
第 4 章 软件可靠性及其测度	53
4.1 软件可靠性的重要意义	53
4.2 软件开发的生命周期	55
4.2.1 启动和结束阶段	56
4.2.2 需求条件和规格说明	56
4.2.3 建立原型样本	57
4.2.4 设计	58
4.2.5 编程	59
4.2.6 测试	60
4.3 软件可靠性及其测度	62
4.4 软件错误及其对软件可靠性模型的影响	63
4.4.1 软件错误与排错曲线	64
4.4.2 软件错误与排错模型	66
4.5 软件可靠性模型	70
4.5.1 常数排错率的软件可靠性模型	70
4.5.2 线性递减型排错率的软件可靠性模型	73
4.5.3 指数递减型排错率的软件可靠性模型	75
4.6 软件可靠性模型中常数估算	77
4.6.1 常数型排错率的参数估算方法	78

4.6.2 线性递减型排错率的参数估算方法	79
4.6.3 指数递减型排错率的参数估算方法	80
4.7 小结	80
4.8 思考题	81
参考文献	82
第5章 冗余技术及其应用	85
5.1 功能性冗余	85
5.1.1 静态功能性冗余	85
5.1.2 动态功能性冗余	88
5.2 结构性冗余	89
5.2.1 软硬件系统的结构性冗余	89
5.2.2 主动冗余	90
5.2.3 被动冗余	95
5.2.4 混合冗余	100
5.2.5 时间冗余	110
5.3 编码技术及其应用	111
5.3.1 检错编码和纠错编码的基本原理	112
5.3.2 线性分组码	116
5.3.3 汉明纠一检二编码	120
5.3.4 萧纠一检二编码	124
5.4 小结	126
5.5 思考题	126
参考文献	128
第6章 避错技术	132
6.1 规格说明阶段应用的避错技术	133
6.1.1 确信技术	134
6.1.2 验证技术	135
6.2 在设计阶段应用的避错技术	138
6.2.1 故障预防——设计过程中的确信技术	139
6.2.2 故障检测——设计过程中的验证技术	141
6.3 设计阶段应用的功能测试	147
6.4 小结	149
6.5 思考题	149
参考文献	149
第7章 防错技术	151
7.1 故障防止应注意的事项	151

7.2 防止故障的措施	152
7.2.1 硬件系统采取的防止故障措施	152
7.2.2 软件系统采取的防止故障措施	154
7.3 小结	157
7.4 思考题	157
参考文献	157
第8章 硬件系统排错技术	158
8.1 测试的基本概念	158
8.1.1 产品测试	159
8.1.2 维护性测试	160
8.2 逻辑测试的基本概念	165
8.2.1 逻辑测试的基本类型	165
8.2.2 逻辑测试中测试生成的基本参数	167
8.2.3 逻辑测试的基本分类	167
8.2.4 逻辑测试的实施步骤	169
8.3 组合电路的测试和测试生成	169
8.3.1 布尔差分	170
8.3.2 D 算法	178
8.3.3 九值算法	186
8.3.4 PODEM 算法	188
8.3.5 可控性和可观察性	191
8.3.6 FAN 算法	194
8.3.7 面向测试码的 Poage 算法及压缩算法	197
8.4 时序逻辑电路的测试生成	202
8.4.1 时序电路测试的基本概念	203
8.4.2 状态表验证和 I/O 校验序列	204
8.4.3 利用鉴别序列生成的校验序列	206
8.4.4 无鉴别序列时序电路的校验序列	211
8.5 桥接故障模型及其测试生成	214
8.5.1 桥接故障的类型	214
8.5.2 非反馈型桥接故障的测试生成方法	215
8.5.3 反馈型桥接故障的测试生成	216
8.6 小结	219
8.7 思考题	219
参考文献	222
第9章 软件系统排错技术	225
9.1 软件测试的故障覆盖率	225

9.2	语句测试	227
9.3	分支和路径测试	228
9.4	条件和决策测试	229
9.5	变异测试法	230
9.5.1	变异测试基本概念	230
9.5.2	变异测试的实施	231
9.6	软件内建自测试的基本思想及实施框架	233
9.7	常用的测试用例生成方法简介	234
9.7.1	黑盒测试用例生成方法	234
9.7.2	白盒测试的测试用例生成方法	234
9.7.3	二叉化基本路径测试的分析法	235
9.7.4	基本路径二叉化的路径覆盖率分析	239
9.8	两种值得注意的软件故障模型的分析	239
9.8.1	变量未初始化	240
9.8.2	空指针	240
9.9	小结	242
9.10	思考题	242
	参考文献	243
第 10 章	可测性设计	246
10.1	可测性设计思想的提出	246
10.2	可测性设计的基本原理	248
10.2.1	测试质量和可测性属性	248
10.2.2	可测性的属性	248
10.3	随机应变法	249
10.3.1	设置附加测试点	249
10.3.2	便于初始化设置	251
10.3.3	将大规模组合电路进行分解的方法	252
10.3.4	提高对时序电路的可控性	253
10.3.5	软件系统测试点设置和异常检测技术	255
10.4	专用可测性电路及可测性软件设计方法	258
10.4.1	Reed-Muller 电路扩展技术	258
10.4.2	控制逻辑插入技术	261
10.4.3	专用可测性设计在软件中的应用	263
10.5	组合电路 BIT 设计方法	264
10.5.1	PLA 电路的结构及基本故障模型	265
10.5.2	PLA 电路的可测性设计	271
10.6	时序电路 BIT 设计方法	274
10.6.1	扫描通路设计思想	274

10.6.2	隔离(切换)部件的设计	275
10.6.3	电平触发扫描设计	277
10.6.4	扫描设计对系统开发成本的影响	279
10.7	边界扫描 BIT 技术	280
10.7.1	边界扫描问题的提出	280
10.7.2	边界扫描设计的基本原理	280
10.8	BIST 方法	283
10.8.1	BIST 的基本概念	283
10.8.2	线性反馈移位寄存器的基本理论	284
10.8.3	一个可测性设计实例	288
10.9	小结	291
10.10	思考题	292
	参考文献	294
第 11 章	容错设计技术	297
11.1	结构性冗余技术	300
11.1.1	软件的 N 版本(模)冗余的基本概念	300
11.1.2	软件的 N 版本冗余技术的实现	302
11.2	卷回和向后恢复技术	303
11.2.1	向后恢复技术	303
11.2.2	向后恢复技术中的高速缓存	303
11.2.3	向后恢复技术中恢复点的确定	304
11.2.4	向后恢复技术中运行环境的恢复	305
11.3	向前恢复技术	306
11.3.1	恢复模块式	306
11.3.2	终结模式技术	307
11.4	各种容错技术的比较	307
11.4.1	容错设计技术的相似性	307
11.4.2	容错设计技术的相异性	308
11.5	运用容错技术与系统可信性的关系	310
11.6	小结	310
11.7	思考题	311
	参考文献	311
第 12 章	故障安全技术	313
12.1	故障安全系统	313
12.1.1	利用固有的安全性	314
12.1.2	利用结构冗余技术的安全性	314
12.1.3	应用结构冗余设计故障安全系统实例	315

12.2 自校验技术	318
12.2.1 双轨校验电路	319
12.2.2 n 取 m 码自校验电路	320
12.3 故障安全系统与自校验系统的比较	322
12.4 小结	323
12.5 思考题	323
参考文献	324

第 1 章 数字计算系统可信性及属性

1.1 可信性定义及问题的提出

近年来,随着信息技术领域的迅速发展,人们加速了对信息技术新领域、新产品(数字计算系统)的研究和开拓。由于信息技术学科理论、技术和产品,不仅广泛应用于诸如航空航天、军事、国防及与国民经济相关的各高科技领域,而且已经发展到与人类社会、人们生活,乃至生命、财产的安全等问题息息相关。因此,人们对信息技术领域里的新技术、新产品必然会提出更高的需求和更严格的条件,即除了在传统意义上要求产品具有更高性能、更低成本,以及便于使用、易于维修等一般性规格和标准外,人们还一致认为更重要的标准则是数字计算系统所提供的服务是否是“可信的”。于是,数字计算系统的可信性(dependability, 也称为诚信性 trustworthy)问题的重要性和迫切性开始备受人们的关注。由于人类的生活乃至生存空间对信息技术及其相关产品的依赖程度愈来愈高,人们很自然地会对任何一个所应用的数字计算系统(或一件产品),提出一个“它能否为自己提供可信服务?”的问题。因为数字计算系统已经不是一般意义上的商品了,它的可信服务已经成为衡量其质量最重要的标准之一。在过去,由于受到各种条件(包括经济、理论和技术)的限制,“可信系统”的研究和开发只是数字计算系统领域中的阳春白雪,对一般用户来说,往往是可望而不可即,仅仅极少数几个发达国家在某些特定要求下,才有可能和能力去开发和应用可信计算系统。比如在航天飞机飞行姿态控制系统、原子核反应堆控制系统、导弹防御系统、铁路运输信号控制系统以及金融部门的监督系统等都需要提供完全可信的服务才可能完成其所承担的任务。这些控制系统中的任何隐患都将给人们带来严重的甚至是灾难性的后果。尤其是在当今的普适计算(pervasive computing)时代中,由于计算系统已经渗透到社会的各个领域,其中包括计算机网络的高度发展和广泛使用,网格计算的日趋普及以及计算系统在人类各种经济活动和社会领域中的应用等,使得人们愈来愈期待着自己所应用和依靠的各种计算系统都是可信赖的。因此,可信计算和可信计算系统的开发和研究必将成为今后计算系统发展的主要趋势。

一个数字计算系统的可信性是指该系统可以提供确实可信服务的综合能力^[1, 2]。之所以称它为一种综合能力,是因为衡量这个能力的标准并不是,也不可能是单一的,而是一个由多方面综合因素确定的评价体系。它主要涉及到如下 6 个属性^[3]: 可靠性(reliability)、可用性(availability)、可测性(testability)、可维护性(maintainability)、安全性(safety)以及保密性(security)等。上述诸方面组成了度量可信性的属性。其中,可靠性是对系统在一定条件下能正常完成所规定任务能力的评价,它的评价标准主要在于系统保持正常运行时间的长度;可测性则是对一个系统具有避免故障、防止故障(即在系统开发阶段,在系统正式投入运行之前和系统发生故障之前就必须考虑到的措施)和检测、诊断故障的能力以及在系统正常运行时,修复故障能力的评价;可用性和可维护性是对系统在具有一定的维护措施

以后,随时可提供正常服务能力的评价;安全性则是对系统一旦发生失效(失效的定义将在1.2节给出)后限制其发生严重后果能力的评价。而保密性乃是关于对系统保持其内部所存信息的信任度和完整性(系统是否被非法入侵、复制和篡改等)能力的评价。本书将对可信性的上述各类属性作详细的分析研究。由于对保密性的评价体系涉及到信息保存、处理、安全、人为因素以及其他方面的知识,因此不属于本书的讨论范围。

随着信息和计算技术的不断发展,近年来,可信性计算理论和技术正在深入地应用于各类计算系统、网络以及网格计算中^[4]。可信计算系统被逐步广泛应用,主要有以下几个原因:

(1) 在恶劣环境中的应用 近年来,由于高集成度的微处理器的发展,计算机系统已经从条件优越的计算机机房走到了工业、农业、国防、通信及交通等领域的现场。现场可能根本没有空调冷却设备,而且温度和湿度随时有较大幅度的变化;震动频繁,电源、电网电压变化甚大以及电磁场的干扰作用等。因此系统必须具有抗恶劣环境的功能——需要可靠性、可用性和可维护性技术。

(2) 在网络环境中的应用 由于有线和无线网络的出现,随之而来的是病毒(virus)、黑客(hacker)以及各种非法入侵行为的破坏和干扰,尤其是对软件系统(包括程序和数据)的攻击和破坏(包括入侵、篡改和扰乱等)使得一些数字系统往往不能提供可信的服务而使正常用户蒙受巨大损失和严重的后果。因而,一个可信计算系统必须具有抗干扰的功能——需要保密性和安全性技术。

(3) 各种不同层次新用户的使用 由于计算机系统的不断更新换代,对一般用户来说经常遇到新的系统(包括新的软件和硬件系统),他们对于一个新系统的性能及操作使用方法往往知道甚少,以致有误用、乱用之处。计算系统也应有允许用户误用的功能——需要可靠性和可维护性技术。

(4) 系统维修成本的不断提高 由于当前的计算机发展有硬件成本不断下降、而劳动力成本(包括软件开发和程序编写工作)不断增高的趋势,使用硬件冗余而减少维修次数的方法不失为一种既合乎发展潮流又具有经济实效的措施。当然硬件的冗余与劳动力成本之间的关系也要掌握一定的比例——需要可测性和可维护性技术。

(5) 在高性能计算及复杂系统中的应用 由于计算机系统愈来愈庞大、复杂,使用的元器件个数愈来愈多,软件规模愈来愈大,使整个大系统发生错误的概率也就提高了。因此,有必要使系统在具有高性能的同时还需要有一定的检测、纠错、修复故障的容错能力,使一个完整的系统不至于由于某一个或几个部件的故障无法正常运行而提供错误的服务——需要可测性、可维护性、可靠性和安全性技术。

本书将从数字计算系统(包括系统的软件和硬件在内)的可信性及可信计算的角度出发,对可信数字计算系统进行分析和论证,同时也提出可信计算系统的设计原理及方法。

1.2 影响可信性计算的主要因素

每一个计算系统在其运行期间都可能受到来自各方面的干扰和影响,以致降低或损害了系统的可信性。其中,影响可信性计算最重要的有如下三个相关的因素:故障(fault)、错误(error)和失效(failure)。事实上,这三个因素是造成系统可信性受损的直接原因,而且三

者之间有着密切的因果关系,下面将分别给出它们的定义,并讨论它们的性质、属性、相互关系和对可信性的影响等问题。

1.2.1 失效的定义及后果

按照事物发展的一般规律,任何一个系统在运行期间都可能出现非正常现象。即,系统在运行到一定的时间,或在一定的条件下偏离它预期设计的要求或规定的功能。这种现象通常称为失效。换句话说,就一般情况而言,系统的失效是不可避免的,是绝对的。而保持系统的正常运行则只是局部的,相对的。同时,由于人们对数字计算系统的依赖程度愈来愈大,因此,一旦一个系统出现失效,给人类带来的后果也将愈来愈严重。例如,美国科罗拉多州首府丹佛的国际机场,由于行李管理系统的失效而延期开张达数月之久。雅丽安娜5号火箭由于电子控制系统中错写了一个常数而引起的失效不得不使它的首次发射日期多次改期。而发射到火星的探测器由于设计中使用了不统一长度单位(英寸和厘米)而导致多次失效。更有甚者,还有不少由于治疗癌症的仪器中控制系统的失效导致多名病人死亡的例子。这些例子都说明了一个事实:功能强大的计算系统在给人类带来福音的同时也存在着失效后产生严重后果的可能性。

进一步来看,失效可以从以下4个不同的角度(称为失效模式)进行分析和讨论:失效的值域范围、失效的时域范围、失效的感受范围和失效后果的严重性等。

(1) 从失效的值域模式来看,有两种情况:

- 数值失效 也称为静态失效,指的是系统提供了非正常(过大或过小)的数据。
- 时间失效 也称为动态失效,指的是系统的响应时间不准确(太快或太慢)。

(2) 失效的时域模式分析,包括两种情况:

- 持续性失效 指系统在一段重要运行期间出现的非正常现象。
- 瞬时性失效 指系统在某一时刻出现短暂的错误行为。

(3) 从失效的感受模式来看,也可分为两种情况:

- 一致性失效 系统所有用户所感觉到的失效是相同的。
- 非一致性失效 不同的用户感觉到的失效情况是不同的。

(4) 从失效后果的严重性方面来分析,有3种可能性:

- 轻微(良性)损坏(失效) 因失效而产生的后果或造成的损失不大于该系统在正常运行时其本身所提供服务的量值。
- 严重损坏(失效) 失效造成一定的经济损失和人员受伤等后果。
- 危险性(恶性)损坏(失效) 因失效而产生的后果或造成的损失大于该系统在正常运行时所提供服务的价值。
- 灾难性损坏(失效) 系统可能发生“崩溃”,并且在系统发生崩溃后,其执行的一切任务也立即停止,其后果是不可接受的。

另外,一个系统失效的后果达到良性失效中最严重的地步,则称为达到了(良性)失效的临界点(criticality)。

从对数字计算系统的失效及其后果分析来看,失效的严重性已经到了非关注不可的地步了。众所周知,一方面由于先进的信息技术为人类社会开创了新时代,给人们带来了不可估量的发展前景,另一方面,它们的失效却给人类造成难以形容的苦楚。事实上,这是一对

难以回避的矛盾和严峻的事实。因此,我们必须面对现实,正视这对矛盾。而要解决这对矛盾,就须要避免或减少由于失效而产生的各种不良后果。为此,首先必须对失效的起因有一个深入的了解。

1.2.2 故障的定义及性质

如上所述,失效是一个系统在运行时发生偏离预期规定而出现非正常现象的一种表现。而产生这种失效现象的根本原因则可以归纳为“故障”。故障这个词在一般日常生活中用得很多。在硬件系统中它既有“缺陷”(defect),“瑕疵”等含义,而在软件技术中也有所谓“隐患”,“病虫”(经常称为 bug)等的意思。尽管它在常规的意义下有一定的模糊性,在本书中还必须对故障作一个明确的定义。事实上,故障可以认为是(系统)产生失效的似然条件和推理上的原因。之所以说它是“推理上的原因”是因为要确切地诊断或识别系统中的故障是一件十分复杂的过程。它涉及到相关的诊断者自身掌握知识水准和信息量的多少,对被诊断对象的功能的理解程度,以及对该系统的控制和观察能力等诸多因素。故障和失效存在着密切的因果关系。然而,从失效现象逆向找寻产生失效的根源——故障,却并没有直接的明显的方法。因此,人们只能根据失效的表现形式,依靠分析和推理,去发现失效的原因。所以,即使找到了一个故障,也只能作为产生失效的一个推理上的原因。另外,即使找到了一个可能的故障,进而需要在系统中确定这个故障的具体位置,也是一个十分复杂的推理过程。例如,对一个硬件系统(产品)中存在的故障,或许只需找到故障所在的模块(如集成电路芯片)即可(称为系统级诊断)。但是,有时还必须作进一步诊断以确定故障在芯片中的哪一个门电路上(称为门级电路诊断),或者更为详细地需要了解故障是由哪一个 MOS 电路或晶体管造成的(称为电路级诊断)等。除此之外,从原因上来分析,引起失效的原因有可能并非是由单个故障造成的,而是由多个故障(包括人为因素等)共同的影响产生的。这样,推导失效的原因就更为复杂了。类似地,对软件系统的故障诊断同样需要考虑到不同模块(或子程序)之间的关系。然后通过对这些模块之间关系进行分析将故障诊断逐步细化,最终来确定故障可能的位置(可能是在某一个模块中,也可能确定到某一行程序,或者是某个符号或参数等)。另外,由于一个软件系统运行必然同它所使用的操作系统的环境和其他同时运行的软件系统有着密切的关联,而这些因素同时也增加了软件系统故障诊断的复杂性。综上所述,得到的结论是,根据对失效的表现,通过分析研究寻找到失效的根源——故障,只能是该失效推理上的起因。

既然检测和诊断故障是一件十分复杂的事情,那就必须从故障的性质和起源这两方面对它作深入的研究。首先,必须弄清楚的是故障的起因、起源:

- 故障来自何处 故障的发源地在哪里?
- 故障何时生成 在系统的生命周期中何时引入了故障?

其次,应该知道故障的基本性质:

- 从故障类型来看 是功能性的,还是技术性的(具体定义将在以后给出)?
- 从形成故障的故意程度来看 造成故障的原因纯粹是偶然(如元器件的老化引起的,还是人为故意(如对网络的非法入侵和信息的篡改)造成的?
- 从故障持续时间来看 是永久性(permanent)故障还是暂时性(temporary)故障?

1. 故障的起源

(1) 故障的发源地 计算系统的故障主要来自于两方面：系统内部故障和系统外部故障。

系统内部故障是指系统本身由于某种原因(包括人为的、非人为的)而产生的故障。它可以在系统的规格说明、设计、生产和运行等各个阶段中产生并显示它的后果。如,由于在飞行器的导航程序中存在某个故障,使得该飞行器在作某些飞行动作时发生失效。又如,在一个数字电子线路中的由于元件的老化而产生**短路故障**(也称为**桥接故障**(bridging fault),本书对短路故障和桥接故障两术语是通用的),该故障有可能导致整个线路的错误动作而引起系统的失效。其中,第一个例子中的故障是在该软件系统的规格说明或是程序设计阶段产生的,而第二个例子的故障则是在系统运行中发生的。

系统外部故障主要来自于人为因素(如输入错误、按键错误、数据错误等)和外部环境的影响(如温度、湿度、震动等),这些原因都可能引起系统中某一部分的故障。

(2) 故障形成的时间 失效一般是在系统的运行中表现出来的,但是,引起失效的故障却可以在系统的整个生命周期都可以形成,包括系统的规格说明阶段、设计阶段、生产制造阶段和运行期间。这是因为在系统生命的每一个阶段都有可能产生各种故障。如,在规格说明时,对系统的功能描述不完全,造成设计者和用户之间的不统一而产生故障。在其他几个阶段中就更不可避免地引起故障的发生。

2. 故障的性质

(1) 从故障的基本性质分析,故障有功能性故障(简称功能故障)和技术性故障(简称技术故障):

- 功能性故障 以人为造成的为主。由于人为(主要是指非故意行为)的原因,在系统整个生命周期都可能引入功能故障。如,在输入程序时,漏输一行语句,或错写一个符号,或者在装配流水线上将逻辑门插错位置,如与门变成了或门等都属于功能故障。由于故障以人为因素为主,这类故障的发生往往有比较确切的原因。在软件系统中主要以这类故障为主。
- 技术(工艺)性故障 又称为物理故障或硬故障,它主要发生在硬件系统。如信号线的开路或短路,导通的半导体突然截止等现象就是这类故障的典型例子。这类故障可以在系统的运行中随时发生。这是因为技术故障受环境和元器件老化的影响很大,因此,它的形成有极大的随机性。虽然,技术故障主要是对硬件系统而言,软件系统中往往很少出现此类故障,但是,随着计算技术的发展,在软件系统中同样可能出现这类故障。这是因为当前操作系统的发展和更新很快,而有一些过时的应用软件不适应在更新过的操作系统环境下运行,因此很可能产生上述的技术故障。

(2) 从引起故障的故意性来分析,有非故意性故障和故意性故障:

- 非故意性故障 它是最常发生的,也是本书研究的主要对象。如,设计者在设计时对需求分析研究不够,甚至于理解错误等,引起规格说明或者设计上的某些故障。又如,程序设计者对程序设计语言语义的错误理解等都可能引起故障。
- 故意性故障 明显的是通过人为的一种手段故意造成的故障,它可能使部分甚至整