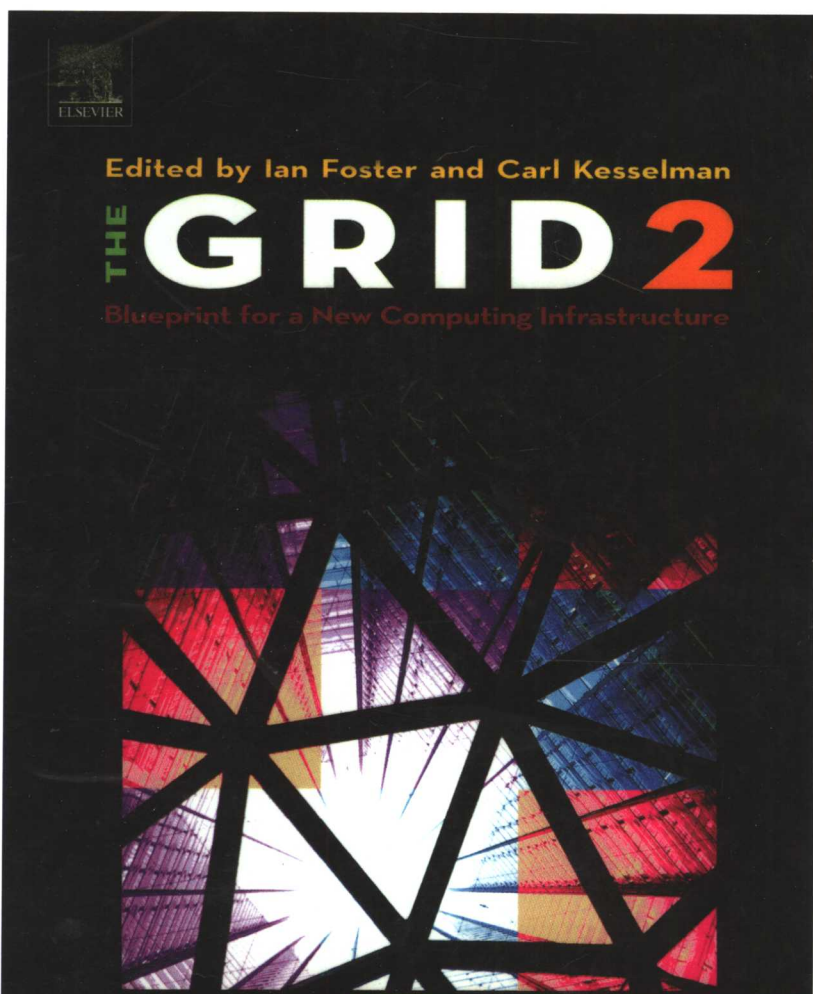


# 网格计算

(英文版 · 第2版)



(美) Ian Foster Carl Kesselman 编著  
芝加哥大学 南加利福尼亚大学

TP301.6  
Y14

经典原版书库

# 网格计算

(英文版·第2版)

The Grid  
Blueprint for a New Computing Infrastructure  
(Second Edition)

江苏工业学院图书馆  
藏书章

(美) Ian Foster Carl Kesselman 编著  
芝加哥大学 南加利福尼亚大学

北方工业大学图书馆



00732867



机械工业出版社  
China Machine Press

05/02

The Grid: Blueprint for a New Computing Infrastructure, Second Edition by Ian Foster and Carl Kesselman (ISBN 1-55860-933-4).

Original English language edition copyright © 2004 by Elsevier Inc. All rights reserved.

Authorized English language reprint edition published by the Proprietor.

ISBN: 981-2593-02-0

Copyright © 2005 by Elsevier (Singapore) Pte Ltd., 3 Killiney Road, #08-01 Winsland House I, Singapore 239519.

Printed in China by China Machine Press under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR and Taiwan.

Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书英文影印版由Elsevier (Singapore) Pte Ltd. 授权机械工业出版社在中国大陆境内独家发行。本版仅限在中国境内（不包括香港特别行政区及台湾地区）出版及标价销售。未经许可之出口，视为违反著作权法，将受法律之制裁。

版权所有，侵权必究。

本书法律顾问 北京市展达律师事务所

本书版权登记号：图字：01-2005-1311

#### 图书在版编目（CIP）数据

网格计算（英文版·第2版）/（美）福斯特（Foster, I.）等编著. —北京：机械工业出版社，2005.4

（经典原版书库）

书名原文：The Grid: Blueprint for a New Computing Infrastructure, Second Edition

ISBN 7-111-16054-1

I. 网… II. 福… III. 网格—计算—英文 IV. O243

中国版本图书馆CIP数据核字（2005）第006971号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：迟振春

北京昌平奔腾印刷厂印刷·新华书店北京发行所发行

2005年4月第1版第1次印刷

787mm×1092mm 1/16·48.5印张

印数：0 001-3 000册

定价：88.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换  
本社购书热线：(010) 68326294

# 出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域中取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅肇划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战，而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall, Addison-Wesley, McGraw-Hill, Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum, Stroustrup, Kernighan, Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及收藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师提供服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国

家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业的教学度身订造的。其中许多教材均已为M. I. T., Stanford, U.C. Berkeley, C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程，而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下，读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证，但我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：

电子邮件：[hzedu@hzbook.com](mailto:hzedu@hzbook.com)

联系电话：(010) 68995264

联系地址：北京市西城区百万庄南街1号

邮政编码：100037

# 专家指导委员会

(按姓氏笔画顺序)

尤晋元  
石教英  
张立昂  
邵维忠  
周立柱  
范明  
袁崇义  
谢希仁

王珊  
吕建  
李伟琴  
陆丽娜  
周克定  
郑国梁  
高传善  
裘宗燕

冯博琴  
孙玉芳  
李师贤  
陆鑫达  
周傲英  
施伯乐  
梅宏  
戴葵

史忠植  
吴世忠  
李建中  
陈向群  
孟小峰  
钟玉琢  
程旭

史美林  
吴时霖  
杨冬青  
周伯生  
岳丽华  
唐世渭  
程时端

## 秘书组

武卫东

温莉芳

刘江

杨海玲

---

# Dedication

For Alexander and Imogen, who have already enjoyed a lifetime of the Grid—Ian

To Phyllis for enduring the process, and Squirrel for keeping me company in the early morning hours—Carl

---

# Editors

Dr. Ian Foster was born in Wellington, New Zealand. He has a bachelor of science (Hons I) degree in computer science from the University of Canterbury in Christchurch, New Zealand, and a doctorate in computer science from Imperial College, London. Foster is Associate Division Director and Senior Scientist in the Mathematics and Computer Science Division of Argonne National Laboratory, and Professor of Computer Science at the University of Chicago. Foster's research interests are in innovative technologies and applications of large-scale distributed computing. An internationally recognized researcher and leader in the area of Grid computing, he has published five books and over 200 articles and technical reports in parallel computing, distributed systems, and advanced applications. Foster is a fellow of the American Association for the Advancement of Science and of the British Computer Society, which recognized his work on the Strand parallel programming language with its 1989 award for technical innovation.

Dr. Carl Kesselman is the Director of the Center for Grid Technologies at the University of Southern California's Information Sciences Institute and an Information Sciences Institute Fellow. He is also a Research Associate Professor of Computer Science at the University of Southern California. He received a Ph.D. in Computer Science from the University of California at Los Angeles, a Master's of Science in Electrical Engineering from the University of Southern California, and Bachelors in Electrical Engineering and Computer Science from the University of Buffalo. Kesselman's research interests are in all aspects of Grid computing, and his research contributions in this area are internationally recognized. He has published over 100 articles and papers in the area of high performance and distributed computing. He is a fellow of the British Computer Society.



Foster and Kesselman have worked together since the late 1980s, collaborating first on projects in parallel computing and more recently within the context of the Globus Project<sup>®</sup>, which they established with their colleague Steven Tuecke in 1996. One significant product of their joint work has been the Globus Toolkit<sup>®</sup>, the open source software that has emerged as the *de facto* standard for Grid computing in both e-business and e-science. Their contributions have been recognized by numerous awards, including the 1997 Global Information Infrastructure Next Generation Internet award, a 2002 R&D 100 award, the 2002 R&D Editor's choice award, and the 2002 Ada Lovelace Medal from the British Computing Society.

They have recently begun to edit a new series entitled *The Elsevier Series in Grid Computing* that will explore all aspects of Grids from standards to applications. Designed for researchers and professionals, the books will be authored by the foremost experts in the field. This second edition of *The Grid* is the first title in this new series.

---

# Preface to the Second Edition

The *Grid* is an emerging infrastructure that will fundamentally change the way we think about—and use—computing. The word *Grid* is used by analogy with the electric power grid, which provides pervasive access to electricity and, like the computer and a small number of other advances, has had a dramatic impact on human capabilities and society. Many believe that by allowing all components of our information technology infrastructure—computational capabilities, databases, sensors, and people—to be shared flexibly as true collaborative tools, the Grid will have a similar transforming effect, allowing new classes of applications to emerge.

This is the second edition of *The Grid: Blueprint for a New Computing Infrastructure*. Writing this preface brings home to us just how much has happened since the first edition was published, how much has been learned, and how much new information there is to communicate about Grid technologies, their application, and likely future evolution.

When we started planning the first edition early in 1997, we believed that Grids were destined to be important, but we would not have predicted the impact that Grid technologies have had on the practice of science and industry in the five years since the book appeared in mid-1998. Rereading the first edition, we are delighted with how much of what we wrote back then still applies. However, it is certainly also the case that much has been learned and much is new.

Our understanding has evolved tremendously with respect to the nature both of the opportunities offered by Grid computing and of the technologies needed to realize those opportunities. We now see far more clearly the revolutionary impacts of large-scale resource sharing and virtualization within both science and industry, the intimate relationships between organizational and resource-sharing structures, and the new technologies required to enable secure, reliable, and efficient resource sharing on a large scale.

This understanding has emerged from aggressive and imaginative application experiments conducted by talented research and development teams around the world. We report on some highlights of this work here, and these stories have much to say concerning the excitement that Grid technologies generate within application communities, the successes that have been achieved, and (we must not lose sight of this perspective) the obstacles that remain to be overcome.

The technology has evolved, too, in both its implementation and its specification. The open source Globus Toolkit® was only just emerging five years ago, when the first edition appeared, and it has now been widely adopted as a de facto standard infrastructure with a substantial international contributor and user base: truly the “Linux of the Grid.” Meanwhile, work within the Global Grid Forum has produced specifications defining the Open Grid Services Architecture, a Grid standard that is seeing broad adoption within industry and research, with implementations available from multiple sources, including the Globus Toolkit. Other complementary technologies—such as Condor, portal tools, and various commercial offerings—have matured to the point where they form vital parts of the overall Grid software ecosystem. Thus, we are able to describe not only what might be but what is, and the architectural framework within which future developments will occur.

Equally remarkable is the flowering of the Grid community. The Global Grid Forum now regularly welcomes close to a thousand participants at its triennial meetings to discuss and define the technical standards required for further evolution of the Grid. Large international science Grid projects in astronomy, physics, biology, and other fields are well under way, and hundreds of smaller projects are developing Grid technologies, deploying Grid infrastructures, and applying Grids to a wide range of challenging applications. In Asia, Europe, and North America, major government initiatives are being established to create the Grid infrastructures required to support twenty-first-century science and engineering. Meanwhile, a growing number of both large companies and startups (several featured in this book) are investing large sums in the technology and, increasingly, selling Grid solutions to companies with challenging information technology requirements.

---

## Goals of This Book

Our goal in the second edition of *The Grid* is to present an up-to-date view of Grids that reflects this progress and that both reports on real experiences and explains the technologies available today and emerging from labs, companies, and standards bodies. Achieving this goal without at least doubling the size of the book

would have been impossible if it were not for the miracle of the Web, and our enlightened editor at Morgan Kaufmann. Fortunately, for your backs and wallets, we have been able to place much of the material from the first edition on the Web, allowing us to focus the second on what is new.

The development of this new edition has thus been an ambitious undertaking, involving substantial revisions to some chapters and a large number of entirely new contributions. In fact, there can be few sentences that remain the same from first edition to second. Nevertheless, the old content remains accessible on the Web, and we refer to it frequently here, using the notation “Chapter 1:N” to refer to Chapter N in the first edition.

This edition, like the first edition, is intended to serve simultaneously as a manifesto, design blueprint, user guide, and research agenda for future Grid systems.

As a *manifesto*, the book seeks to make the case for Grids, explaining why they are important and why we should be concerned with making them a reality.

As a *design blueprint*, the book is intended for software engineers, computing center managers, technology program managers, chief information officers, and other professionals who will be tasked with developing and deploying Grids. This blueprint is intended to describe what Grids will look like, how they will be constructed, and how they will operate.

As a *user guide*, the book is intended for scientists, engineers, and programmers who will develop applications for Grids. For these users, we want to explain what Grid applications will look like, how they will be developed, and what new ways of thinking and programming they will require.

As a *research agenda*, the book is intended for students, researchers, and research managers who want to understand the state of the art in relevant disciplines and the areas in which gaps in our knowledge demand further research. We hope thus to broaden the engagement of researchers in distributed systems, security, collaborative work, and other relevant domains whose engagement is vital to the long-term success of the Grid agenda.

Clearly, we cannot expect to do justice to all four goals in a single book. However, we do believe that we have succeeded in taking useful steps toward each goal, and in so doing we hope to advance understanding of the technologies needed to create the Grids of the future and the implications of these technologies for future applications of computers.

The construction of Grids draws upon expertise from many diverse branches of computer science, including distributed systems, networking, computer and network security, electronic commerce, information systems, databases, operating systems, high-performance computing, and algorithms. Hence, in preparing this book we have assembled a distinguished set of authors, each recognized as an authority in one or more of these fields. We have asked these authors to

summarize concrete achievements and the state of the art in their particular technology area, indicating where further developments will likely take place and the achievements expected within the next five to ten years. We have also asked them to identify the primary obstacles to progress—the areas in which focused research efforts can be expected to make a significant difference. The result, we hope, is a book that, like the first edition, will be valuable for years to come, both as a summary of where we are and as a road map for where we need to go to make Grids a reality.

---

## Structure and Content of This Book

As summarized in Table 1 we have divided this second edition into seven parts comprising a total of 30 chapters. We lead off in Part I with three personal statements by influential thinkers on the significance of Grids, from the perspectives of infrastructure, science, and industry, respectively. This material provides context for Part II, which comprises a single scene-setting chapter that lays out central Grid concepts and architectural principles upon which the rest of the material in the book is based. Part III is concerned with applications. Building on the first edition's discussion of compute-intensive, data-intensive, teleinstrumentation, and collaborative applications (Chapters 1:3–6, respectively), this edition's 12 application vignettes each present a substantial application success story. Parts IV, V, and VI expand upon the architectural presentation of Part II by providing detailed discussions of core architecture and services, data and knowledge management, and higher-level tools, and Part VII addresses production Grid deployment, the computing platforms on which Grids are constructed, peer-to-peer technologies, and network infrastructure.

As this brief preview indicates, this book covers a tremendous amount of material, more than we expect most readers to handle in one sitting. Thus, we have structured chapters so that each can be read in isolation, while also providing (through careful editing and numerous cross references) the structure that ensures that the whole is indeed more than the sum of its parts.

---

## Text Use

The book is designed to be used by the practicing professional or as a text for a senior undergraduate- or graduate-level course in advanced networking, distributed computing, or Grids. In a quarter-length course, one lecture can be spent on

Section	Topic	Chapter
Perspectives	Grids in Context	1
	The Scientific Imperative	2
	The Industrial Imperative	3
Framework	Concepts and Architecture	4
Applications (brief "vignettes")	Predictive Maintenance: Distributed Aircraft Engine Diagnostics	5
	Distributed Telepresence: The NEESgrid Earthquake Engineering Collaboratory	6
	Scientific Data Federation: The World-Wide Telescope	7
	Medical Data Federation: The Biomedical Informatics Research Network	8
	Knowledge Integration: In Silico Experiments in Bioinformatics	9
	Distributed Data Analysis: Federated Computing for High-Energy Physics	10
	Massively Distributed Computing: Virtual Screening on Desktop Computers	11
	Enterprise Resource Management: Applications in Research and Industry	12
	Interactivity with Scalability: Infrastructure for Multiplayer Games	13
	Service Virtualization: Infrastructure and Applications	14
	Group-Oriented Collaboration: The Access Grid Collaboration System	15
	Collaborative Science: Astrophysics Requirements and Experiences	16
Architecture	The Open Grid Services Architecture	17
	Resource and Service Management	18
	Building Reliable Clients and Services	19
	Instrumentation and Monitoring	20
	Security for Virtual Organizations: Federating Trust and Policy Domains	21
Data and Knowledge	Data Access, Integration, and Management	22
	Enhancing Services and Applications with Knowledge and Semantics	23
Tools	Application-Level Tools	24
	Languages, Compilers, and Runtime Systems	25
	Application Tuning and Adaptation	26
Infrastructure	Production Deployment: Experiences and Recommendations	27
	Computing Elements	28
	Peer-to-Peer Technologies	29
	Network Infrastructure	30

each chapter, and a project may be assigned based on one of the Grid programming systems described in the book. In a semester-length course, some topics can be covered in greater depth. Each chapter includes recommendations for further reading.

In addition, we maintain a Web site that provides pointers to relevant Web resources, including online documents and information about projects described in the book. The URL for this Web site is [www.mkp.com/grid2](http://www.mkp.com/grid2). Questions or comments about this book or the Web site should be directed to the book's general editors at [Gridbook@globus.org](mailto:Gridbook@globus.org).

---

## Acknowledgments

It is with great pleasure that we acknowledge the efforts of the many people who have contributed to the development of this book. First and foremost, we thank the contributors, who despite busy schedules devoted much effort to planning the book, writing their chapters, and responding to numerous comments and suggestions from the editors and reviewers. In addition to the distinguished authors from the first edition, most of whom also contribute here, we welcome new contributors from many countries (the Grid is truly an international undertaking) and from industry.

We also thank the many colleagues and supporters who have contributed to the progress that we report in this second edition. Here, our acknowledgments must unfortunately be far from complete. However, we must certainly recognize the visionary and dedicated government program managers who funded early research and development work that has made Grid computing possible: notably, Mary Anne Scott and Tom Kitchens at the U.S. Department of Energy Office of Science; Gary Koob at DARPA; Alan Blatecky, Frederica Darema, Dick Hilderbrandt, and Rich Hirsh at the National Science Foundation; Kyriakos Baxevanidis at the European Union; and Tony Hey and John Taylor in the UK eScience programme. We also thank the pioneers who have adopted, applied, promoted, or otherwise advanced the work: David Abramson, Malcolm Atkinson, Paul Avery, Fran Berman, Charlie Catlett, Tom DeFanti, Bill Feireisen, Fabrizio Gagliardi, Lennart Johnsson, Bill Johnston, Ken Kennedy, Miron Livny, Satoshi Matsuoka, Mirco Mazzucato, Paul Messina, Jeff Nick, Dan Reed, Ed Seidel, Satoshi Sekiguchi, Larry Smarr, Rick Stevens, and Irving Wladawsky-Berger, to name just a few. And last but not least, this work would not have been possible without the support of our colleagues at Argonne National Laboratory, the University of Chicago, and the University of Southern California's Information Sciences Institute.

Finally, we gratefully acknowledge the expert help we received during the preparation of the manuscript. Gail Pieper of Argonne National Laboratory once again provided invaluable editing assistance, working far beyond the call of duty to enable us to submit a high-quality manuscript. Julie Wulf-Knoerzer of Argonne handled many aspects of the manuscript production process in her usual efficient manner. Also, as was the case with the first edition, we have been fortunate to work closely with an outstanding team at Morgan Kaufmann Publishers. Denise Penrose, Emilia Thiuri, Alyson Day, Angela Dooley, and the rest of the team at Morgan Kaufmann Publishers were everything we could want from a publisher: professional, efficient, and a delight to work with.

Ian Foster and Carl Kesselman



---

# Contents

Editors	vii
Preface to the Second Edition	ix

---

## **PART I PERSPECTIVES** **1**

---

### **1 Grids in Context** **3**

*Larry Smarr*

- 1.1 Lessons from History 3
- 1.2 The Grid on Steroids 6
- 1.3 From Networks to the Grid 7
- 1.4 Who Will Use the Grid? 8
- 1.5 Summary 12

---

### **2 The Scientific Imperative** **13**

*Fran Berman and Tony Hey*

- 2.1 New Modes of Scientific Inquiry 15
- 2.2 Virtual Community Science 19
- 2.3 Investing in New Infrastructure 21
- 2.4 Summary 23