



# Grid Computing

## THE SAVVY MANAGER'S GUIDE

Pawel Plaszczyk and Richard Wellner, Jr.

# **Grid Computing**

# The Savvy Manager's Guides

*Series Editor, Douglas K. Barry*

*Grid Computing*

Pawel Plaszczak

Richard Wellner, Jr.

*Semantics in Business Systems*

Dave McComb

*Business Intelligence*

David Loshin

*Web Services and Service-Oriented  
Architecture*

Douglas K. Barry

# Grid Computing

## The Savvy Manager's Guide

---



Pawel Plaszczak  
Richard Wellner, Jr.



AMSTERDAM • BOSTON • HEIDELBERG • LONDON  
NEW YORK • OXFORD • PARIS • SAN DIEGO  
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Morgan Kaufmann Publishers is an imprint of Elsevier



MORGAN KAUFMANN PUBLISHERS

Senior Editor	Rick Adams
Publishing Services Manager	Simon Crump
Project Manager	Brandy Lilly
Assistant Editor	Mona Buehler, Rachel Roumeliotis
Cover Design	Francis Bacca
Cover Image	"Teotitlán" by Pawel Plaszczak
Text Design	Julio Esperas
Composition	SPI Publisher Services
Technical Illustration	Dartmouth Publishing, Inc.
Copyeditor	SPI Publisher Services
Proofreader	SPI Publisher Services
Indexer	SPI Publisher Services
Interior printer	The Maple-Vail Book Manufacturing Group
Cover printer	Phoenix Color

Morgan Kaufmann Publishers is an imprint of Elsevier.  
500 Sansome Street, Suite 400, San Francisco, CA 94111

This book is printed on acid-free paper.

© 2006 by Elsevier Inc. All rights reserved.

Designations used by companies to distinguish their products are often claimed as trademarks or registered trademarks. In all instances in which Morgan Kaufmann Publishers is aware of a claim, the product names appear in initial capital or all capital letters. Readers, however, should contact the appropriate companies for more complete information regarding trademarks and registration.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, scanning, or otherwise—without prior written permission of the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333, e-mail: [permissions@elsevier.co.uk](mailto:permissions@elsevier.co.uk). You may also complete your request on-line via the Elsevier homepage (<http://elsevier.com>) by selecting "Customer Support" and then "Obtaining Permissions."

#### Library of Congress Cataloging-in-Publication Data

Plaszczak, Pawel.

Grid computing : the savvy manager's guide / Pawel Plaszczak, Richard Wellner, Jr.  
p. cm.

Includes bibliographical references and index.

ISBN 0-12-742503-9 (alk. paper)

1. Computational grids (Computer systems) I. Wellner, Richard. II. Title.

QA76.9.C58P65 2005

004¢.36—dc22

2005012671

ISBN-13: 978-0-1274-2503-0

ISBN-10: 0-12-742503-9

For information on all Morgan Kaufmann publications,  
visit our Web site at [www.mkp.com](http://www.mkp.com) or [www.books.elsevier.com](http://www.books.elsevier.com)

Printed in the United States of America

05 06 07 08 09 5 4 3 2 1

Working together to grow  
libraries in developing countries

[www.elsevier.com](http://www.elsevier.com) | [www.bookaid.org](http://www.bookaid.org) | [www.sabre.org](http://www.sabre.org)

ELSEVIER

BOOK AID  
International

Sabre Foundation

# Contents

Foreword . . . . .	xiii
Preface . . . . .	xv
1. Introduction . . . . .	1
Problems of today and promise of Grid computing. . . . .	5
A vision of the Grid. . . . .	6
Example 1: Techman runs simulations using on-demand CPU-hours . . . . .	6
Example 2: HydroTech offers secure, remote processing of sensitive data. . . . .	9
Example 3: InvestIT automatically acquires information on foreign markets for niche customers . . . . .	10
The feel of the Grid today . . . . .	12
Google and the World Wide Web. . . . .	13
Travelocity, Sabre, and Global Distribution Systems . . . . .	17
Summary . . . . .	23
2. The Basics . . . . .	25
This chapter in two paragraphs . . . . .	25
Scientific roots . . . . .	26
Early proto-grids (meta-computers) . . . . .	29
Noncomputational resources . . . . .	33
Virtualization . . . . .	35
Note on the term <i>virtualization</i> . . . . .	37
Scientific roots: Summary. . . . .	37
Business perspective. . . . .	38
Standing on the shoulders of giants . . . . .	38
Web services . . . . .	39
Service-oriented architecture. . . . .	39
Today's Web services are not enough. . . . .	40
The prospect of utility computing . . . . .	41

Business meets academia . . . . .	42
WS-Resource Framework . . . . .	44
Note on the term <i>Grid service</i> . . . . .	45
The meaning of WSRF . . . . .	45
Note: OGSF, the predecessor to WSRF . . . . .	46
Virtual organization . . . . .	47
Security for virtual organizations. . . . .	48
The case for Open Grid Service Architecture. . . . .	50
OGSA services: Overview. . . . .	53
Virtual organization . . . . .	53
Composition, grouping, orchestration, workflow . . . . .	53
Transactions . . . . .	54
Metering, accounting, and billing . . . . .	54
Installation, deployment, provisioning . . . . .	54
Application contents . . . . .	54
Information and monitoring . . . . .	55
Logging . . . . .	55
Messaging. . . . .	55
Security . . . . .	55
Policy . . . . .	56
Data. . . . .	56
Program execution . . . . .	56
Grid computing defined and redefined . . . . .	57
Note on spelling of “grids,” “Grid,” and “Grid computing” . . . . .	60
Disagreement on the definitions . . . . .	60
Cluster of clusters . . . . .	61
Cycle scavenging. . . . .	61
Grid versus distributed computing. . . . .	61
Grid versus Web services . . . . .	62
Grid versus peer-to-peer (P2P). . . . .	62
Other definitions of Grid computing. . . . .	63
Will there be the Grid? . . . . .	63
 3. Grid Computing Enters Business. . . . .	 65
How the market understands grids . . . . .	66
Departmental grids. . . . .	66
Enterprise grids. . . . .	68

Partner grids . . . . .	69
Open grids and the Grid . . . . .	70
Taxonomy summary . . . . .	71
What is the market? . . . . .	72
The educational and public research market . . . . .	72
Engineering and manufacturing . . . . .	73
Life sciences . . . . .	73
Financial sector . . . . .	74
The future market . . . . .	74
How can market segmentation hamper open grids? . . . . .	74
Case study: The MCNC Enterprise Grid . . . . .	76
Service . . . . .	77
Customers . . . . .	78
Financials . . . . .	79
Resources . . . . .	80
Location . . . . .	81
Who is who: The ecosystem of the Grid . . . . .	82
Standard bodies and consortia . . . . .	83
Academic projects and groups of interest . . . . .	84
Large vendors . . . . .	85
Grid-specializing vendors and niche vendors . . . . .	86
Application vendors . . . . .	87
Vendor taxonomy according to the 451 . . . . .	88
Grid resource providers . . . . .	88
Consultants . . . . .	89
Media and other information providers . . . . .	90
The panorama of today's products . . . . .	90
Distributed resource management . . . . .	91
SUN N1 Grid Engine family . . . . .	91
LSF Suite from Platform Computing . . . . .	92
Service containers and hosting environments . . . . .	94
Note on service containers, server containers, and service provisioning systems . . . . .	95
Globus Toolkit 4 Grid service container . . . . .	95
Portals and end-user interfaces to the Grid . . . . .	97
Security solutions . . . . .	98
Data solutions . . . . .	99
Avaki . . . . .	99
Applications . . . . .	101

Case Study: The NEESgrid cyberinfrastructure . . . . .	101
Technical overview . . . . .	103
Hardware . . . . .	105
Software . . . . .	106
Planning the communication with users. . . . .	106
Testing and quality management . . . . .	108
Accomplishments . . . . .	109
Summary . . . . .	109
 4. Joining the Grid . . . . .	 111
Strategies for participating. . . . .	113
What does it mean to “plug in”? . . . .	114
Build an enterprise grid. . . . .	114
Case study: Wachovia’s Grid-powered trade floor . . . . .	116
Business case. . . . .	117
Transition . . . . .	117
Applications . . . . .	118
Standards . . . . .	119
Summary . . . . .	119
Participate in a partner grid. . . . .	120
Join the industry-wide Grid as end-user . . . . .	123
Case study: Synopsys: Software release engineering on the Grid . . . .	125
Deploy monitoring before transition . . . . .	127
The transition. . . . .	127
Technology. . . . .	129
Problems . . . . .	130
Summary . . . . .	131
What does it mean to provide a Grid-enabled solution? . . . . .	131
Become a Grid infrastructure provider . . . . .	131
Become a service provider on the Grid. . . . .	134
Provide a Grid-enabled product. . . . .	136
Refactor application for horizontal scalability and on-demand business model . . . . .	136
Refactor applications for Grid-compliant connectivity . . . . .	138
Become a middleware vendor . . . . .	140
Case study: Rolls-Royce and Data Systems & Solutions develop grid for Equipment Health Monitoring . . . . .	141
The support services model . . . . .	141
Data systems & solutions . . . . .	142

Equipment Health Monitoring (EHM) problem description. .	143
DAME. . . . .	146
Generalization . . . . .	146
Analysis . . . . .	147
Summary . . . . .	149
 5. Technical Issues. . . . .	 151
High-level system design. . . . .	151
Organizational security requirements and firewalls . . . . .	152
Data sensitivity. . . . .	154
CPU peak requirements and duty cycles . . . . .	155
Data storage . . . . .	156
Internet bandwidth availability . . . . .	157
Existing resources . . . . .	159
Custom resources . . . . .	160
Potential for porting . . . . .	162
Side trip: Analogies. . . . .	163
The Web . . . . .	163
Peer to Peer . . . . .	164
CORBA, DCOM. . . . .	164
Technology Areas . . . . .	165
Data management and databases. . . . .	165
Secure remote access . . . . .	166
Taking advantage of clusters . . . . .	167
Secure delegation . . . . .	167
Federated databases. . . . .	167
Storage management . . . . .	169
Resource management . . . . .	173
Supercomputers . . . . .	174
Clusters and farms . . . . .	175
On-demand CPU resources. . . . .	175
Workflow management. . . . .	176
Guaranteed execution . . . . .	177
Security . . . . .	177
Internal versus external security . . . . .	178
Techman case revisited . . . . .	179
Feasibility study . . . . .	182
Authorization scalability and federations . . . . .	183
Other security concerns. . . . .	185
Pragmatic view. . . . .	185

Other Grid security solutions . . . . .	186
Where do the standards go? . . . . .	188
Security summary . . . . .	189
 6. Management Issues . . . . .	 191
Building and selling Grid business case . . . . .	191
Identify problems to solve . . . . .	193
Marketing mix . . . . .	196
Product and product differential advantage . . . . .	196
Price . . . . .	198
Promotion . . . . .	201
Place (channels) . . . . .	202
Press . . . . .	203
Events . . . . .	204
Partnering . . . . .	205
SWOT matrix . . . . .	206
Strengths . . . . .	206
Weaknesses . . . . .	206
Opportunities . . . . .	207
Risks . . . . .	207
Change and transition period management . . . . .	207
Change driven by technology . . . . .	208
Change driven by organization . . . . .	208
Change resulting from introduction of new products . . . . .	209
Change readiness . . . . .	209
Identify charismatic leaders . . . . .	209
Motivate people . . . . .	210
Minimize hierarchy . . . . .	210
Planning the change . . . . .	211
Engage team to identifying solutions . . . . .	211
Construct the vision shared by everyone . . . . .	212
Build a sensible deployment plan . . . . .	212
Start transition from peripheries and let it radiate . . . . .	213
Adjust strategy to the observed transition process . . . . .	213
Address human factor . . . . .	213
Single change versus constant change . . . . .	214
Migrate in three stages . . . . .	214

Scale up the pilot system . . . . .	215
Internal change first, external later . . . . .	215
Role of consultants . . . . .	215
Do you need a consultant? . . . . .	216
Precisely Define the consultant's mission . . . . .	217
Understand the interest of the consultant . . . . .	217
Be actively involved in the mission. . . . .	217
Engage external resources for diagnosis. . . . .	218
Gradually transfer responsibilities. . . . .	218
Be prepared to take over . . . . .	218
Risk mitigation. . . . .	218
Execution risk management: the traditional approach . . . . .	219
Risk identification . . . . .	219
Risk quantification . . . . .	220
Risk response development . . . . .	220
Risk response control. . . . .	220
Fighting white space risk through rapid-result initiatives . . . . .	221
Result-oriented . . . . .	222
Vertical . . . . .	222
Fast . . . . .	223
Independent . . . . .	223
White space in the Grids sector . . . . .	223
Focus on selected elements of SOA . . . . .	224
Run the complete system for a short while . . . . .	224
Build complete microscale system . . . . .	225
Agile development . . . . .	225
Anticipate shift of requirements . . . . .	227
Refactor rather than design in advance . . . . .	228
Engage developers in decision making . . . . .	228
Code is the documentation . . . . .	229
Pair programming . . . . .	230
Test-driven programming. . . . .	231
Institutionalize the process. . . . .	232
The Globus Campaign System. . . . .	232
Treat customers as partners . . . . .	233
Treat developers as partners . . . . .	234
Summary. . . . .	235
Afterword . . . . .	237
Index. . . . .	239



# Foreword

Like many disruptive technologies before it, such as the Internet and Web, the Grid was initially motivated by the needs of scientists. Originally designed to facilitate access to high-end computational resources such as supercomputers, the Grid quickly expanded to enable sharing of a wide range of resources including storage, networks, and scientific instruments such as microscopes, x-ray sources, and earthquake engineering test facilities. Common to all of these applications was the need to form and maintain distributed virtual communities within which resources could be pooled and applied to common problems.

As the Grid enters its second decade, its applications are extending far beyond these pioneering scientific users. The need to share resources to achieve common goals turns out to be fundamental to commerce as well as science. Whether to support business processes across partners in a supply chain, facilitate collaborative engineering among subcontractors, or enable higher utilization of resources spread across business units, resource sharing provided by Grid technologies is becoming increasingly critical to a wide range of business and commercial activities.

As the Grid crosses the boundary from elite scientific research into mainstream business, the supporting literature needs to change and support the new audience. So far, the Grid has been described primarily in scientific articles. The few books that have been published on the subject have been written in technical language and directed to an academic audience or to developers ready to get down in the trenches.

For these reasons, this is a timely book: the savvy manager needs to know about the Grid; what makes it tick; and how it can be used in his or her organization to decrease costs, improve performance, and enable business processes that would otherwise be intractable. Pawel Plaszczak and Richard Wellner take on the brave task of addressing this need and provide the first ever comprehensive, yet compact work directed to CEOs, IT managers, planners, and decision makers, and those who are interested in deploying and using the Grid paradigm without necessarily programming it. They explain why Grid is

important and provide the information needed to evaluate whether a transition to Grid is right for you, and how to plan such a move.

The authors of this book are particularly well qualified to make the case for Grid computing to the business IT manager. Both have been intimately involved in developing and deploying Grid technology and applications for many years. Pawel and Richard have both been major contributors to the open source Globus Toolkit, the most widely used Grid infrastructure to date. In addition to their deep knowledge of the fundamental technology, they have both also been part of the commercial sector, developing applications of Grid technology with major corporations.

In this book, Pawel and Richard provide the background knowledge that the savvy manager needs to understand the potential of the Grid for his or her organization. A rich assortment of case studies illustrates the range of problems that the Grid is being used to solve and will aid the reader in identifying which of their problems are “Grid” problems. These case studies also show how to get started and introduce the steps that should be followed to incorporate the Grid into existing IT infrastructures.

The impact of the Grid is already being felt, as corporations roll out large deployments and ISVs start to Grid-enable their products. However, “you ain’t seen nothing yet”: the next ten years will be exciting indeed and seem likely to lead to radical changes in the way that we interact with our IT infrastructure. With this book, you, the “savvy manager,” can understand how to gain benefits today and how to be prepared for the future. We think that you will find it an interesting and useful read.

Ian Foster and Carl Kesselman