

# **Grid Application Systems Design**

**April J. Wells**



**Auerbach Publications**

Taylor & Francis Group  
Boca Raton New York

---

Auerbach Publications is an imprint of the  
Taylor & Francis Group, an **informa** business

---

# Contents

---

<b>Preface</b> .....	<b>xiii</b>
<b>Acknowledgments</b> .....	<b>xix</b>

## **PART I: IN THE BEGINNING**

<b>1 History</b> .....	<b>3</b>
Computing .....	3
Methodology .....	4
Paradigm .....	4
Data Directed Programming .....	5
Procedural Programming .....	5
C .....	7
C++ (C plus plus) .....	7
BASIC (Beginner's All-Purpose Symbolic Instruction Code) .....	8
COBOL .....	9
FORTRAN .....	10
Perl .....	10
Structured Programming .....	11
Ada .....	12
Pascal .....	12
Unstructured Programming .....	13
Imperative Programming .....	15
ALGOL .....	16
Modula 2 .....	17
Modula 3 .....	18
Oberon .....	19
Declarative Programming .....	19
Functional Programming .....	20
Flow-Driven Programming .....	20

Event-Driven Programming .....	20
Class-Based Programming .....	21
Java .....	21
C# .....	22
Prototype-Based Programming .....	23
Component-Oriented Programming .....	23
Object Linking and Embedding (OLE) .....	25
Component Object Model (COM) .....	25
Concept-Oriented Programming.....	26
Programming Paradigms and Grid .....	26
<b>2 Definition and Components.....</b>	<b>29</b>
Definition .....	29
Peer-to-Peer (P2P) Computing.....	31
Napster.....	31
Gnutella .....	32
Types .....	33
Computational Grid.....	33
Distributed Servers and Computation Sites .....	34
Remote Instrumentation.....	34
Data Archives .....	35
Networks.....	36
Portal (User Interface) .....	36
Security .....	37
Brokers .....	37
User Profiles.....	38
Searching for Resources.....	38
Batch Job Submittal.....	39
Credential Repository .....	40
Scheduler .....	41
Data Management.....	41
Data Grid.....	42
Storage Mechanism Neutrality .....	42
Policy Neutrality.....	43
Compatibility with Other Grid Infrastructures .....	43
Storage Systems .....	43
Access or Collaboration Grids .....	44
Large Format Displays.....	44
Presentation Environments.....	44
Interfaces to Grid Middleware.....	44
Other Components.....	45
Scavenging Grid .....	45

Grid Scope.....47  
 Project Grid, Departmental Grid, or Cluster Grid .....47  
 Enterprise Grid or Campus Grid.....47  
 Global Grid.....47

**3 Early Adopters.....49**

Computational and Experimental Scientists .....49  
 Bioinformatics.....50  
 Corporations.....50  
 Academia.....50  
 University of Houston, Texas, United States. ....50  
 University of Ulm, Germany.....51  
 The White Rose University Consortium, United Kingdom.....51  
 Science.....52  
 Particle Physics.....52  
 Industry.....53  
 Gaming.....53  
 Financial .....54  
 Wachovia.....55  
 RBC Insurance.....55  
 Charles Schwab .....55  
 Life Sciences .....56  
 American Diabetes Association .....56  
 North Carolina Genomics and Bioinformatics Consortium.....57  
 Spain’s Institute of Cancer Research.....58  
 Petroleum .....58  
 Royal Dutch Shell.....58  
 Utilities.....58  
 Kansai Electric Power Co., Inc.....58  
 Manufacturing.....58  
 Ford Motor Company .....58  
 Saab Automobile .....59  
 Motorola Semiconductor.....59  
 Project MegaGrid ..... 60  
 Southern Partnership for Advanced Computational  
 Infrastructure (SPACI).....61  
 Chicago Stock Exchange.....62  
 Ohio Savings Bank .....62  
 Governmental Agencies .....62  
 NASA.....63  
 U.S. Department of Defense .....63

European Union.....	64
Flemish Government.....	64
Benefits.....	65
Virtualization.....	65

## PART II: THE PARTS AND PIECES

<b>4 Security.....</b>	<b>71</b>
Security.....	71
Authentication .....	71
Passwords and Personal Identification Numbers.....	72
Public Key Infrastructure and Digital Certificates.....	73
Tokens .....	73
Biometrics.....	75
Geography.....	76
Passwords .....	78
Private Key Cryptography.....	78
Block Ciphers .....	80
Stream Ciphers .....	80
Public Key Cryptography .....	80
Digital Signature .....	84
Authorization.....	88
Delegation of Authority.....	88
Accounting.....	89
Audit .....	89
Access Control.....	89
Usage Control.....	92
Cryptography.....	92
Block Cipher.....	93
Stream Ciphers .....	93
Data Integrity.....	97
Capability Resource Management .....	100
Database Security.....	102
Inference.....	102
<b>5 Hardware.....</b>	<b>115</b>
Computers .....	115
Blade Servers .....	120
Storage.....	121
I/O Subsystems.....	123
Underlying Network.....	124

Operating Systems .....	124
Visualization Environments .....	124
People .....	125
<b>6 Metadata.....</b>	<b>127</b>
Defining Metadata.....	127
Grid Metadata.....	131
Data Metadata .....	132
Physical Metadata.....	133
Domain-Independent Metadata .....	133
Content-Dependent Metadata .....	133
Content-Independent Metadata.....	134
Domain-Specific Metadata .....	134
Ontology.....	134
User Metadata .....	134
Application Metadata .....	134
External Metadata .....	135
Logical Metadata .....	135
User .....	136
Data.....	136
Resources.....	136
Metadata Services .....	136
Context .....	136
Content.....	137
Structure .....	137
Defining Data Granularity.....	137
XML.....	137
Database.....	139
Access .....	139
Metadata Formatting.....	139
XML .....	140
What is XML?.....	140
Application .....	144
MCAT .....	145
Conclusion.....	146
<b>7 Drivers .....</b>	<b>147</b>
Business .....	149
Accelerated Time to Results .....	149
Operational Flexibility .....	150
Leverage Existing Capital Investments .....	151
Better Resource Utilization .....	152

Enhanced Productivity.....	152
Better Collaboration.....	153
Scalability.....	154
Return on Investment (ROI).....	155
Reallocation of Resources.....	156
Total Cost of Ownership (TCO).....	157
Technology .....	158
Infrastructure Optimization.....	158
Increase Access to Data and Collaboration.....	158
Resilient, Highly Available Infrastructure .....	159
Most Efficient Use of Resources .....	159
Services-Oriented Approach.....	161
Batch-Oriented Approach .....	161
Object-Oriented Approach.....	162
Supply and Demand.....	162
Open Standards .....	162
Corporate IT Spending Budgets.....	163
Cost, Complexity, and Opportunity .....	163
Better, Stronger, Faster .....	164
Efficiency Initiatives .....	165

## PART III: APPLICATIONS IN THE GRID

<b>8 Virtualization.....</b>	<b>169</b>
Definition .....	172
Why Virtualization Matters .....	172
How to Virtualize .....	175
Problems .....	176
ABI/API Emulation .....	177
<b>9 Strategy and Planning.....</b>	<b>183</b>
Introduction .....	183
Design .....	186
Inputs and Outputs .....	186
Parallelism versus Fragmentation .....	187
Data-Based Parallelism.....	189
Processing .....	190
Data Storage.....	190
Signoff.....	191
Implementation.....	191
Launching Solution.....	192

Technical Documentation .....	194
User Documentation .....	194
Evaluation and Maintenance.....	195
Services .....	196
Analysis of a System.....	197
Types of Systems .....	198
Investigations .....	201
<b>10 Programming .....</b>	<b>205</b>
Grid Programming .....	205
Parallel Programming.....	205
Multi-Threaded Programming.....	212
Web Services.....	213
SOAP .....	214
WSDL.....	214
WS-Inspection .....	215
<b>11 Wrapping It Up.....</b>	<b>217</b>
<b>Index.....</b>	<b>219</b>