ORACLE®

Oracle Press™

Oracle Database 11*g* Oracle Real Application Clusters Handbook

Second Edition

K Gopalakrishnan

McGraw-Hill

New York Chicago San Francisco
Lisbon London Madrid Mexico City
Milan New Delhi San Juan
Seoul Singapore Sydney Toronto

Contents

Foreword xiii
Acknowledgments xvii
Introduction xix

PART I
High Availability Architecture and Clusters

1 Introduction to High Availability and Scalability 3
High Availability 4
HA Terminology 5
Planned and Unplanned Outages 5
An End-to-End Perspective 6
Cost of Downtime 6
Building Redundant Components 8
Common Solutions for HA 10
Cluster, Cold Failover, and Hot Failover 10
HA Option Pros and Cons 14
Scalability 14
Oracle Real Application Clusters Solution 14
Emerging Trends 15
Oracle 11*g* Solutions 16
In a Nutshell 16

2 Clustering Basics and History 17
Grid Computing with Clusters 19
Shared Storage in Clustering 23
Types of Clustering Architectures 23
History of Oracle RAC 28
Oracle Parallel Storage Evaluation 29
Oracle Parallel Server Architecture 31

Components of an OPS Database 32
Cluster Group Services (CGS) 32
Distributed Lock Manager (DLM) 33
Locking Concepts in Oracle Parallel Server 33
Cache Fusion Stage 1, CR Server 37
Limitations of Oracle Parallel Server 38
The Oracle RAC Solution 39
Availability 40
Scalability 40
Affordability 40
In a Nutshell 41

3 Oracle RAC Architecture 43
Single-Instance vs. Oracle RAC Environment 45
Oracle RAC Components 46
Shared Disk System 47
Oracle Clusterware 53
Oracle Clusterware Components 53
Networking Stack Components 64
Oracle Kernel Components 68
Global Cache and Global Enqueue Services 68
Global Resource Directory 68
Oracle RAC Background Processes 69
In a Nutshell 74

PART II
Installation, Configuration, and Storage

4 Oracle Grid Infrastructure Installation 77
Preinstallation Tasks 79
Configuring the Network 80
Setting Up the Groups and Users 81
Configuring Shared Storage 82
Secure Shell and User Limits Configuration 85
Configuring the Kernel Parameters 86
Oracle Validated Configuration RPM 86
Running the Cluster Verification Utility 87
Oracle Grid Infrastructure Installation 91
Installing the Oracle Grid Infrastructure 94
In a Nutshell 112

5 Oracle RAC Installation 113
Oracle Real Application Clusters Installation 115
Creating the Oracle RAC Database 125
In a Nutshell 135

6 Automatic Storage Management 137
Facts about Automatic Storage Management 138
Physical Limits of ASM 139
ASM in Operation 140
ASM Building Blocks 142
ASM Administration and Management 147
ASM Instance Management 147
ASM Tools 160
ASMCA: The ASM Configuration Assistant 160
ASMCMD: The ASM Command-Line Utility 161
ASM FTP Utility 161
ASMLib 162
Installing ASMLib 163
Configuring ASMLib 163
In a Nutshell 164

PART III
Oracle RAC Administration and Management

7 Oracle RAC Basic Administration 169
Initialization Parameters 170
Unique Parameters 171
Identical Parameters 174
Instance-Specific Parameters 177
Managing the Parameter File 177
Starting and Stopping Instances 178
Using srvctl to Start/Stop Instance(s) 179
Using SQL*Plus to Start/Stop Instances 179
Registering a Single-Instance Database in OCR Using SRVCTL 180
Administering Undo 180
Automatic Undo Management 180
Manual Undo Management 181
Administering a Temporary Tablespace 181
Administering Online Redo Logs 182
Enabling Archive Logs in the Oracle RAC Environment 183
Enabling the Flashback Area 184
Managing Database Configuration with SRVCTL 185
Managing Database Objects 187
Managing Tablespaces 187
Managing Sequences 188
Managing Tables 188
Managing Indexes 188
Scope of SQL Commands 188
Database Connections 189
In a Nutshell 189

8 Oracle RAC Advanced Administration 191
Understanding Services 192
Service Characteristics 193
Administering Services 196
Creating Services 197
Administering SCAN 199
Administering Cluster Ready Services 200
Clusterware Startup Process (Oracle 11*g* R1) 200
Clusterware Starting Sequence 201
Oracle Clusterware Auto Startup 205
Oracle Clusterware Manual Startup 205
Starting and Stopping CRS (Oracle 11*g* R2) 205
Verifying CRS 206
Disabling and Enabling CRS 206
CRS Utilities 207
Administering OCR 214
Administering Oracle Local Registry 216
Administering the Voting Disk 217
In a Nutshell 217

9 Oracle RAC Backup and Recovery 219
Introduction to Backups 220
Oracle Backup Options 220
Oracle Backup Basics 221
Performing Backups in Oracle 222
Instance Recovery in Oracle RAC 223
Redo Threads and Streams 224
Redo Records and Change Vectors 224
Checkpoints 224
Crash Recovery 226
Steps in Crash Recovery (Single Instance) 226
Crash Recovery in Oracle RAC 227
Instance Recovery 227
Instance Recovery in OPS 227
Instance Recovery in Oracle RAC 228
Crash Recovery and Media Recovery 228
Bounded Recovery 229
Block Written Record (BWR) 229
Past Image (PI) 229
Two-Pass Recovery 230
Cache Fusion Recovery 231
Dynamic Reconfiguration and Affinity Remastering 232
Fast Reconfiguration in Oracle RAC 233
Internals of Cache Fusion Recovery 235

Backup and Recovery of the Voting Disk and OCR 238
Backup and Recovery of Voting Disks 239
Backup and Recovery of OCR 239
In a Nutshell 241

10 Oracle RAC Performance Management 243
Oracle RAC Design Considerations 244
Oracle Design Best Practices 244
Oracle RAC–Specific Design Best Practices 245
Partitioning the Workload 246
Scalability and Performance 247
Choosing the Block Size for an Oracle RAC Database 248
Using Automatic Segment Space Management 248
Limitations of ASSM 249
Introduction to the V$ and GV$ Views 249
V$ Views Containing Cache Fusion Statistics 250
Oracle RAC Wait Events 250
Understanding Cluster Waits 251
Global Cache Statistics 263
Global Cache Statistics Summary 264
Global Cache Service Times 264
Global Cache Service Times Summary 267
Enqueue Tuning in Oracle RAC 267
Oracle AWR Report 269
Interpreting the AWR Report 270
GCS and GES Messaging Statistics Section 274
STATSPACK 276
ADDM 276
Tuning the Cluster Interconnect 277
Verifying That Private Interconnect Is Used 277
Interconnect Latencies 278
Verifying That Network Interconnect Is Not Saturated 279
In a Nutshell 279

PART IV
Advanced Concepts in Oracle RAC

11 Global Resource Directory 283
Resources and Enqueues 284
Grants and Conversions 285
Locks and Enqueues 286
Cache Coherency 287
Global Enqueue Services 288
Latches and Enqueues 288

Global Locks Database and Structure ... 289
Messaging in Oracle RAC ... 293
Global Cache Services ... 296
Lock Modes and Lock Roles ... 296
Consistent Read Processing ... 300
GCS Resource Mastering ... 303
Read-Mostly Locking ... 305
In a Nutshell ... 310

12 A Closer Look at Cache Fusion ... 311
Key Components in Cache Fusion ... 313
Ping ... 313
Deferred Ping ... 314
Past Image (PI) Blocks ... 314
Lock Mastering ... 315
Types of Contention ... 315
Cache Fusion I or Consistent Read Server ... 316
Cache Fusion II or Write/Write Cache Fusion ... 319
Cache Fusion in Operation ... 322
Cache Fusion Walkthrough ... 328
Resource Mastering and Remastering ... 346
Background Process and Cache Fusion ... 347
LMON: Lock Monitor Process ... 347
LMS: Lock Manager Server ... 347
LMD: Lock Manager Daemon Process (LMDn) ... 348
LCKn: Lock Process (LCK0) ... 348
DIAG: Diagnostic Daemon (DIAG) ... 348
In a Nutshell ... 348

13 Workload and Connection Management ... 349
Workload Distribution and Load Balancing ... 350
Hardware and Software Load Balancing ... 352
Load Balancing and Oracle Net Services ... 352
Client-Side Load Balancing ... 352
Server-Side Load Balancing ... 356
Transparent Application Failover ... 358
TAF Considerations ... 359
Workload Management ... 365
Oracle Services ... 366
Workload Balancing ... 368
High Availability Features ... 372
In a Nutshell ... 379

14 Oracle RAC Troubleshooting ... 381
Installation Log Files ... 382

Log Directory Structure in the Oracle RDBMS 383
Log Directory Structure in Oracle Grid Infrastructure 384
Troubleshooting a Failed Oracle Grid Infrastructure Installation 385
Inside the Database Alert Log 388
RAC ON and OFF 391
Database Performance Issues 392
    Hung Database 392
Debugging Node Eviction Issues 395
    Cluster Health Monitor 397
    Instance Membership Recovery 398
Advanced Debugging for Oracle Clusterware Modules 405
Debugging Various Utilities in Oracle RAC 407
In a Nutshell 408

PART V
Deploying Oracle RAC

15 Extending Oracle RAC for Maximum Availability 411
Extended Benefits 413
    Full Utilization of Resources 413
    Extreme Rapid Recovery 413
Design Considerations 414
    Speed of Light 414
    Network Connectivity 414
    Cache Fusion Performance 415
    Data Storage 415
Common Techniques for Data Mirroring 416
    Array-Based Mirroring 416
    Host-Based Mirroring 416
    ASM Preferred Read 419
    Challenges in Extended Clusters 419
Extended Oracle RAC Limitations 420
Extended Oracle RAC vs. Oracle Data Guard 421
In a Nutshell 421

16 Developing Applications for Oracle RAC 423
Application Partitioning 426
    Best Practice: Application Partitioning Schemes 427
Data Partitioning 428
    Best Practice: Guidance Systems 429
Buffer Busy Waits/Block Contention 430
Index Partitioning 432
    Buffer Busy Waits: Index Branch/Leaf Blocks Contention 432
    Sorted Hash Clusters 434

Working with Sequences 435
CACHE and NOORDER 435
CACHE and ORDER 436
NOCACHE and ORDER 436
Best Practice: Use Different Sequences for Each Instance 436
Connection Management 437
Full Table Scans 437
Identifying Full Table Scans 438
Interconnect Protocol 439
Library Cache Effect in the Parsing 440
Commit Frequency 441
In a Nutshell 441

PART VI
Appendixes

A Oracle RAC Reference 445

B Adding and Removing Cluster Nodes 473

C References 479

Index 487