

# Colossal Data Warehouse

by

*Azhar Anwar*



**Registration # 1910**

**Enrollment # 244012-045**

*Supervised by*

**Mr. Nadeem Mehmood Khan**

A report is submitted to the Department of Computer Science,  
Bahria Institute of Management & Computer Sciences, Islamabad

In partial fulfillment of requirement for the Degree of

**Master in Computer Science**  
(Majors in Communication Networks)

**Department of Computer Sciences**

**Bahria Institute of Management & Computer Sciences, Islamabad**

**Bahria University, Islamabad**

**June, 2003**



**In the Name of Allah**  
**The Most Beneficent**  
**The Most Merciful**

“In the creation of the heavens and the earth, and in the alteration of night and day, are surely signs to men of understanding, such as keep before their minds the laws of Allah, standing and sitting and reclining, and reflect on the creation of the heaven and earth saying our Rubb # Thou has not created this in vain”.

(Al-Quran Al-Karim 3:189)



# Dedication

*To the life*



## Acknowledgements

First and foremost, I thank Allah for giving the talent and strength to do this and everything else in my life. All of my efforts on this project work and everything else are for his glory and not my own.

I want to thank Mr. Nadeem Mehmood (Project Supervisor) for putting up with me as I raced to get this project done. Without his support and guidance, it would be really difficult to achieve this task.

My indebtedness and gratitude to Mr. Siddique Khilji who have helped in shaping this project cannot adequately be conveyed in a few sentences.

Many thanks to all the respected teachers and well wishing friends who extended towards whatever the help was needed. I also want to thank my loving parents and real friends who encouraged and supported in studies and whole of our lives.



## Project in Brief

Project Title:	Colossal Data Warehouse (CDW)
Objectives:	Develop a massive, cost-effective, reliable and portable Data Warehouse that has storage optimization on other medium than hard disks and the readiness retrieval of data from that medium
Undertaken By:	Azhar Anwar
Supervised By:	Mr. Nadeem Mehmood Khan
Tools Used:	Oracle 9i Database Management System Enterprise Edition, Oracle Developer Forms 6i and PL/SQL
Operating System:	Windows 2000
System Used:	Intel Pentium III, 600 MHz



# Abstract

*Data Warehouse* is a subject-oriented, integrated, non-volatile, time variant collection of data in support of management's decision. Data Warehouse provides instant access to historical data, which is very useful to forecast the trend that is not possible in mainframes (because they don't have the ability to do real-time analysis). The application CDW is developed, providing sufficient data for allowing decisions to be made in a more timely fashion, enhancing the competitiveness and cost structure of the corporation.

*Colossal Data Warehouse* as its name depict possesses the ability to store enormous amount of data. The key strategy used for the development of CDW is its dual storage that can be made on both *hard disk* and *compact discs*. The potential database chosen and developed for CDW is *ISP Billing System*. This database is developed using *Oracle 9i Database Management System*. Features like portability, reliability, and reduced use of resources in terms of system, bandwidth, end-users and market enhance the vision of this application. A *Distributed Database* approach is a very major constituent of this application as two types of data retrieval can be expected in general. One is for query/transaction processing, which is fulfilled by the use of *Production Database*. The second approach for data retrieval can be for the purpose of analysis that is fulfilled from the *Historical Database*. CDW is a client/server application design i.e. based on two tiers.

The application *Colossal Data Warehouse* is completely a new idea and is designed in such a way as to facilitate end-user browsing and execution of queries, both from a performance perspective and ease of understanding. For this purpose the application's front end is designed using *Developer Forms 6i*. This Project report guides you through the path carved for achieving such an application that requires a lot more than just an effort.

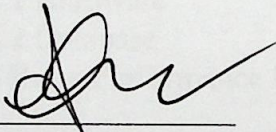


**Bahria Institute of Management & Computer Sciences,  
Islamabad**

**(Department of Computer Science)**

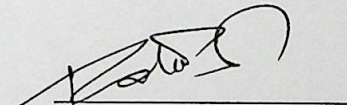
**Certificate**

We accept the work contained in this report as a confirming to the required standard for the partial fulfillment of the degree of MCS (Majors in Communication Networks) in the subject of Computer Science.



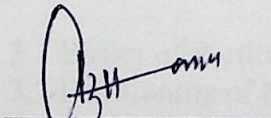
---

Head of Department



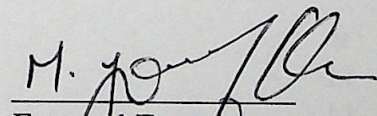
---

Supervisor



---

Internal Examiner



---

External Examiner



# Table of Contents

	Page No.
<b>Chapter 1</b>	
<b>Introduction</b>	1
1.1 Distributed Environment	3
<b>Chapter 2</b>	4
<b>Data Warehouse Requirements</b>	
2.1 Resource Requirements and Dependencies	5
2.1.1 Hardware	5
2.1.2 Database	5
2.2 Basic Performance Enhancements	6
<b>Chapter 3</b>	8
<b>Partitioning</b>	
3.1 Basics of Partitioning	9
3.2 Partitioning of Manageability	11
3.3 Partitioning for Performance	11
3.4 Partitioning for Availability	13
<b>Chapter 4</b>	14
<b>Snapshots</b>	
4.1 Snapshot	15
4.2 Advantages of Using Snapshots	16
4.2.1 Ease Network Loads	16
4.2.2 Mass Deployment	16
4.2.3 Data Subsetting	17
4.2.4 Disconnected Computing	17
4.2.5 Snapshot Architecture	17
4.2.6 View	18
4.2.7 Snapshot Groups	18
4.3 Schema	19
4.4 Database Link	20
4.5 Privileges	21
4.6 SNP Background Processes and Interval	21
4.6.1 Instance Manager	22



## **Chapter 5** 25

### **Configuring Windows 2000**

5.1 Tuning Windows for Oracle	27
5.2 Optimizing Windows 2000 for Oracle	28
5.2.1 Know When Rebooting Is Required	28
5.3 Microsoft Management Console	28
5.3.1 Computer Management Console	28
5.3.2 System Summary	29
5.3.3 Services	29
5.3.4 Disk Management	30
5.3.5 Event Viewer	31
5.3.6 Managing Users and Groups	31
5.3.7 Remotely Managing Servers Using the Computer Management Console	32
5.4 Command Line Management	34
5.5 System Administration	34
5.5.1 Using the Event Viewer	34
5.5.1.1 MMC Integration	35
5.5.1.2 Save Logs in Various Formats	35
5.5.2 Managing Disks	36
5.5.2.1 Online Disk Management	36
5.5.2.2 Remote Disk Administration	36
5.5.2.3 Dynamic Disks and Volumes	36
5.5.2.4 Mount Points	37
5.5.3 The Task Scheduler Service	38
5.6 Remote Administration with the Windows 2000	38
5.6.1 Telnet Server	38
5.7 Resources	40
5.7.1 Microsoft Resources	40
5.7.2 Oracle Resources	41
5.8 Looking Ahead	41

## **Chapter 6** 42

### **Oracle 9i as DBMS**

6.1 Application Development	43
6.1.1 PL/SQL, C and C++ Development	43
6.1.2 Globalization - Unprecedented Unicode Support	44
6.2 Data Integration	44
6.2.1 Capture	45
6.2.2 Staging	46
6.2.3 Consumption	46
6.3 Manageability	47



6.3.1 Simplify Administration	47
6.3.2 Build In Product Intelligence	48
6.3.3 Eliminate Complexity	49
6.3.4 Automate Functionality	49
6.4 Business Intelligence Platform	50
6.4.1 Scalability, Performance, and Manageability	50
6.4.2 Complete business intelligence platform	51
6.4.3 Scalability and Performance	53
6.5 Security	54
6.6 Windows Integration	54
6.7 Oracle9i Database Edition Used	55
6.8 Oracle Partitioning	56
6.9 Oracle Advanced Security	56
6.9.1 Oracle Label Security	56
6.9.2 Oracle OLAP	56
6.10 Feature and Product Availability	57

## **Chapter 7** 64

### **Proposed System**

7.1 Functional Specification	65
7.2 Tools Used	66

## **Chapter 8** 67

### **Software Engineering Using UML**

8.1 Plan and Elaborate Phase	68
8.1.1 Use Case Diagram	68
8.1.2 Describing Processes	68
8.2 Analysis Phase	71
8.2.1 Conceptual Model	71
8.3 System Sequence Diagram	74
8.3.1 Contracts	75
8.4 Design Phase	76
8.4.1 Describing Real Use Cases	76

## **Chapter 9** 79

### **Data Warehouse Model**

9.1 Designing Detail Data	82
9.2 Designing summary data	84
9.3 10 Steps for Designing a Data Warehouse	85
9.4 ER-Diagram	86



9.5 Logical Attributes	87
9.6 Physical Attributes	91
9.7 Database Configuration at “CNT”	91
9.7.1 Database and Instance Information	91
9.7.2 SGA Information	92
9.7.3 Schema	92
9.7.3.1 Schema Objects Summary (non SYS and SYSTEM)	92
9.7.4.1 General User Account Information	92
9.7.4.2 User Roles	92
9.7.4.3 User Tablespace Quota	93
9.7.5 Storage	95
9.7.5.1 Control files	95
9.7.5.2 Tablespaces	95
9.7.5.3 Datafiles	96
9.7.5.4 Rollback Segments	96
9.7.5.5 Redo Logs	96
9.7.6 Schedule	96
9.7.6.1 Initialization Parameters	97
9.7.6.2 Materialized Views	97
9.7.6.3 Refresh Groups	98
9.8 Database Configuration at “POP”	98
9.8.1 Database and Instance Information	98
9.8.2 SGA Information	99
9.8.3 Schema	99
9.8.3.1 Schema Objects Summary (non SYS and SYSTEM)	99
9.8.4 Security	99
9.8.4.1 General User Account Information	99
9.8.4.2 User Roles	100
9.8.4.3 User Tablespace Quota	102
9.8.5 Storage	102
9.8.5.1 Control files	102
9.8.5.2 Tablespaces	102
9.8.5.3 Datafiles	103
9.8.5.4 Rollback Segments	104
9.8.5.5 Redo Logs	104
<b>Chapter 10</b>	<b>105</b>
<b>Future Enhancements</b>	
10.1 List Partitioning	106
10.2 Merge	108
10.3 Multi-table Insert	108
10.4 Data Block Prefetching	109
10.5 Automatic Memory Management in PGA	109



<b>Bibliography</b>	112
<b>Appendix-A</b>	A-1
<b>User Manual</b>	
Application Login Screen	A-2
Welcome Screen	A-3
Insertion in table "Plans"	A-4
Insertion in table "Exchange"	A-5
Insertion in table "Accounts"	A-6
Insertion in Table "Clients"	A-7
Insertion in table "Transaction"	A-8
Insertions in Table "Receipt"	A-9
View Bill	A-10
View Bill Month	A-11
Adding new Month	A-12
View Code tables	A-13
About CDW	A-14
CDW Help	A-15
<b>Appendix "B"</b>	B-1
<b>Programming Code</b>	



## List of Figures

Figure 1: Performance enhancement of Oracle9i for all test results	6
Figure 2: Snapshot Connected to a Single Master Site in Replicated Environment	5
Figure 3: Snapshot Replication Mechanisms	18
Figure 4: Snapshot Groups Correspond with Master Groups	19
Figure 5: Recommended Schema and Database Link Configuration	21
Figure 6: Use Instance Manager to Configure the Number of Job Processes	23
Figure 7: Find the computer name and username	33
Figure 8: Configure Event Log Properties	35
Figure 9: Use Case Diagram	69
Figure 10: Conceptual Model	74
Figure 11: System Sequence Diagram	75
Figure 12: ER-Diagram	87
Figure 13: List Partitioning vs Range Partitioning Graph	109
Figure 14: Memory Management	112
Figure 15: Application Login Screen	A-2
Figure 16: Welcome Screen	A-3
Figure 17: Table Plan insertion screen	A-4
Figure 18: Table Exchange insertion screen	A-5
Figure 19: Table Accounts Screen	A-6
Figure 20: Table Clients insertion screen	A-7
Figure 21: Table Transaction Screen	A-8
Figure 22: Table Receipt Screen	A-9
Figure 23: View Bill Login	A-10
Figure 24: View Bill Month	A-11
Figure 25: Add Month	A-12
Figure 26: Code table view	A-13
Figure 27: About CDW	A-14
Figure 28: CDW Help	A-15



## List of Tables

Table 1: Comparison of Oracle 9i Standard and Enterprise	59
Table 2: Logical Attributes	92
Table 3: Physical Attributes	92
Table 4: Database and Instance Information	92
Table 5: SGA Information	93
Table 6: Schema Objects Summary	93
Table 7: General User Account Information	93
Table 8: User Roles	95
Table 9: User Tablespace Quota	96
Table 10: Control Files	96
Table 11: Tablespaces	96
Table 12: Datafiles	97
Table 13: Rollback Segments	97
Table 14: Redo Logs	97
Table 15: Schedule	98
Table 16: Initialization Parameters	98
Table 17: Materialized Views	99
Table 18: Refresh Groups	99
Table 19: Database and Instance Information	99
Table 20: SGA Information	100
Table 21: Schema Objects and Summary	100
Table 22: General User Account Information	100
Table 23: User Roles	103
Table 24: Datafiles	104
Table 24: Rollback Segments	104
Table 25: Redo Logs	105