

Chapter 1

Discovering Databases and Oracle 12c

In This Chapter

- ▶ Defining databases
 - ▶ Getting to know Oracle databases
 - ▶ Identifying when using Oracle 12c is appropriate
 - ▶ Understanding database administration
-

Welcome to the exciting world of databases, specifically Oracle 12c. Oracle is a recognized hardware and software giant on par with Microsoft in terms of size and influence. Oracle Corporation made its reputation and initial wealth with database technology and is still a recognized leader in databases. Many information technology (IT) professionals would agree Oracle databases provide the most advanced and capable means of storing, accessing, and managing data in the world today. As you read this book, you find out why many professionals hold that belief.

In this chapter, we explain what databases are, how they work, and why they're a core component of any computer system. Next, we introduce Oracle databases and guide you on when to use Oracle 12c. Finally, we examine database administration and what a database administrator's (DBA) role entails.

Discovering Databases

Since the earliest days of *information technology* (that's data processing for you old-timers), data has been at the heart of everything computers do. Users input data; and programs process, analyze, and perform calculations on the data. Then data output occurs. Notice a common theme? Everything relates to data! As technology increasingly becomes a part of daily life, you can expect almost every aspect of your life to have a data element associated with it.

Data comes from different sources in various formats. Every year, new types of data are created as technology evolves. Consider this partial list of data sources:

- ✓ Order information from your favorite online shopping application
- ✓ Trouble ticket information entered into a computer by a person in a call center
- ✓ Financial data including calculations for interest, taxes, and investments generated by banks and investment companies
- ✓ Multidimensional scans captured by complex medical imaging equipment
- ✓ Inventory data captured by tracking products by their RFID tags and point of sale information
- ✓ Images, videos, and audio files uploaded to social media and video sharing websites

The amount of data in the world is growing at an explosive rate. Consider the number of online sales transactions every day in addition to the Twitter tweets and Facebook posts from smartphones. Additionally, almost every retail sale is tied to an automated sales tracking and inventory control system. Almost anything you do that involves interaction, communication, or commerce generates data.

The size and number of data elements created are also increasing as technology evolves. Pictures, sound files, and videos take far more space on computer disk storage than simple text data (letters and numbers). Scientific and medical devices, such as satellite images and MRI scans, are common examples of how large data is generated rapidly and are part of peoples' lives.

Defining a database

All this data needs to reside somewhere; specifically, within a *database*. Databases are complex software programs that catalog and provide access to data. Although the data is usually stored on a disk storage system, the database software manages how that data is stored and managed.

Businesses use databases to store their data because the size and complexity of their data often exceed what can be stored on a spreadsheet. Most business and scientific activities, and many social activities involving technology will require the creation, access, updating, or deleting of data at some level. It is database software that allows these activities occur in a fast, efficient manner.

Databases provide these primary functions:

- ✔ **Storage:** Data must be kept within a computer system, usually disk storage, so that the data is available when needed. Disk storage technologies can vary, but they must be fast and large and are often expensive and complex.
- ✔ **Organization:** Data must be stored in a logical manner on disk so it can be found quickly and efficiently. Compare an organized desk and file cabinet with a child's messy bedroom; where would you prefer to work?
- ✔ **Access:** Finding the requested data in a fast and efficient manner and returning that data to the requestor is a key function of databases.
- ✔ **Security:** A database determines who can access the data and what they can do with that data. Security must be established and enforced.
- ✔ **Adding, updating, and deleting:** After data is added to the database, it can be modified and deleted. The database software manages the complex rules for how data is manipulated.
- ✔ **Safekeeping:** Databases keep data safe and available when failures occur such as a disk drive crashing. Backup and recovery are the processes used by databases to ensure data is not permanently lost if there is a failure.

Here's the general process:

1. You open the database and enter your data.
2. The database determines how best to store and catalog that data in a secure manner on the computer system's disk storage.
3. When you (or a computer application) want to access, update, or delete that data, the database checks your security permissions and then implements your request in the fastest manner based on the rules that have been defined.



Databases can look more like full operating systems (OSs) than simple applications you start and stop. Advanced databases (such as Oracle) have users, detailed security privileges, network configuration settings, performance tuning for memory, CPU, and disk options, and robust backup and disaster recovery (DR) options.

Examining how databases work

Databases have evolved over the years but have stabilized into the *relational model*. A relational database management system (RDBMS) stores data in logical structures called *tables*. A table is a logical container of similar data. A *table definition* is the set of rules or characteristics for each row of data stored in table. When data is loaded into a database, it is loaded as rows within one or more tables based on the characteristics of the data.

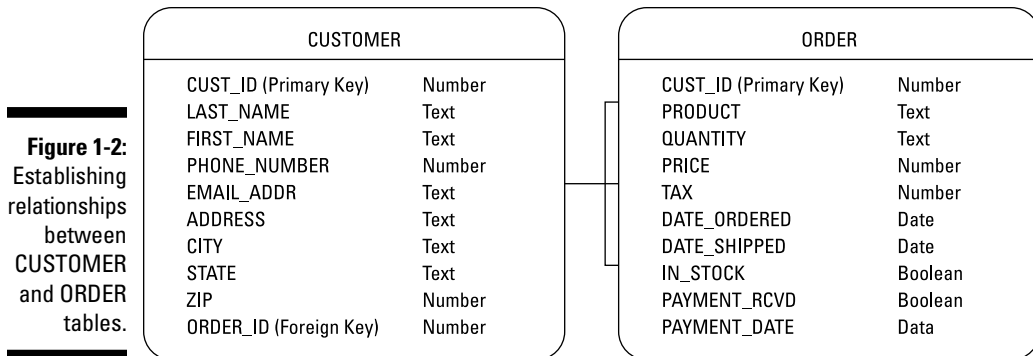
Each row of unique data exists only once in a table. For example, there will only be one row of data for each individual customer in a table. Each row is identified by its *primary key*, which is a unique identifier for that specific row. This uniqueness is a fundamental component of the relational aspect of databases. In Figure 1-1, each customer exists as a single row in the CUSTOMER table as identified by its unique primary key (CUST_ID) and is defined by *columns* specific to that table.

Tables are joined together by connecting the primary key of one table to a related table where it is defined as a *foreign key*. You implement relationships between tables by joining a row of one table to one or more rows in another table. (Remember: This is a relational database).

Figure 1-2 shows how a customer stored in a CUSTOMER table is joined to an ORDER table via the foreign key relationship.

Figure 1-1:
Defining
rows in the
CUSTOMER
table.

CUSTOMER	
CUST_ID (Primary Key)	Number
LAST_NAME	Text
FIRST_NAME	Text
PHONE_NUMBER	Number
EMAIL_ADDR	Text
ADDRESS	Text
CITY	Text
STATE	Text
ZIP	Number
ORDER_ID (Foreign Key)	Number



The CUSTOMER table contains a single row for each customer, and each customer has a unique primary key identifier. The ORDER table stores all the orders for the company and each order is identified by its primary key ORDER_ID. A customer may have zero (new customer), one, or many (repeat customer) orders, and each order must have a customer. (You can't have an order without a customer.) Each row in the ORDER table is tied to the CUSTOMER table by the CUST_ID column which acts as the foreign key joining the ORDER and CUSTOMER table.

Databases have hundreds or even thousands of tables, keys, and relationships, and tables can have millions of rows of data consuming gigabytes or even terabytes of disk storage. To speed access when searching for a specific row of data, an *index* is created on one or more columns in a table. Indexes work similarly to the indexes in a book. On a daily basis, it is common for thousands of rows of data to be inserted, updated, or deleted within the database. The RDBMS software, managed by the database administrator (DBA), supports the overall management, functioning, and performance of the database. We talk about DBAs in the upcoming “Keeping a Database Safe and Sound” section.

Structured Query Language (SQL) is the language used to query (SELECT), create (INSERT), modify (UPDATE), and remove (DELETE) data in a database. SQL is the core language that DBAs and database application developers work in; you need a working knowledge of SQL to effectively manage a database. SQL is further defined as data manipulation language (DML) for querying, creating, modifying, and removing data and data definition language (DDL) for changing database and table structure. Oracle offers a programming extension of SQL called PL/SQL used to implement application logic within an Oracle database.



Oracle gained a big advantage early in the database market by successfully implementing *row locking* inside tables. That is, when one or more people update the same row in a database at the same time, the RDBMS software ensures that everyone's changes are made without conflicts and the data is visible with read consistency throughout the life of the transaction. Not all database vendors handled this issue cleanly.

Finding the right database for the job

The traditional saying, “If all you have is a hammer, everything looks like a nail,” is appropriate to the selection of IT toolsets. Whether people are storing the addresses, phone numbers, and e-mails for a family reunion or they're maintaining an online shopping application, they need a data store. The size, complexity, and scope of that data store determines whether a database is even necessary and, if so, what characteristics (and vendor) of the database are appropriate.

The data requirements for people and businesses vary greatly in size, complexity, and importance. Different technologies exist based on these factors to meet users' data needs.

- ✔ Small, individual, or limited-use data stores, such as address information, are best served by a Microsoft Excel spreadsheet or a Microsoft Access database because both are inexpensive and easy to use.
- ✔ Small or medium size data stores, such as those found at small or medium sized businesses, could make good use of Microsoft SQL Server, Oracle MySQL, or Oracle Enterprise or Standard Edition. These implementations are commonly small, less complex databases where keeping costs low is an important factor.
- ✔ Medium to large data stores, such as those found at medium and large businesses, are commonly the realm of Microsoft SQL Server, IBM DB2, and Oracle Enterprise Edition. This medium to large database market is what many people think of as a typical database environment.
- ✔ Large and extremely large data stores, such as those found at large and multinational businesses, are commonly served by Oracle Enterprise Edition and IBM DB2. The large and complex database environments are special creatures and require very specialized software and hardware.

Review your current computer system environment and projected future needs to make the best decision for your mission. Vendors might try to upsell you on their product, so do your homework before making a decision. However, understand that as your data requirements change, your database environment might require change. We often see systems that started as small, departmental desktop applications grow until the old software is no longer functional. The result is an upgrade to a larger-scaled database system, and that system is often Oracle.

Choosing Oracle

Many seasoned IT people agree that Oracle makes the most advanced and reliable databases in the world. Your authors happen to agree with that opinion, too. Oracle offers several databases targeted for different uses and audiences at varying price points, including Oracle Enterprise Edition, Oracle Standard Edition, Oracle Berkeley DB, Oracle NoSQL, and MySQL. (This book focuses on Oracle Enterprise Edition, Release 12c, although what you learn will be relevant to previous versions of Oracle.) Additionally, Oracle offers a programming extension of SQL called PL/SQL, which is used to implement application logic within an Oracle database.

Oracle as a database

Oracle databases are extremely advanced and sophisticated software components that are state of the art for RDBMS. Many features and options are available, and we guide you through the features you need for most implementations. The highly advanced features, such as Real Application Clusters (RAC), remote data replication, and engineered solutions such as Exadata, are for customers who require the cutting edge of technology for specialized implementations. Visit www.oracle.com for more details on these advanced technologies.

The sophistication and technological advancement that make Oracle databases so great can also make them very complex. Skilled DBAs devote their entire careers to learning the finer points of Oracle databases — it's a never-ending (but often rewarding) process.

Fortunately, Oracle understands that not everyone wants to commit themselves to learning the complexities of Oracle databases. Therefore, Oracle provides a robust, web-based administration tool, Enterprise Manager, where you manage and monitor your databases and the environment supporting them. Leveraging Enterprise Manager to make your job easier is a theme you can find throughout this book.

What is the “c” in Oracle 12c?

Following on the success of the previous Oracle database, release 11g (*g* is for *grid*), the most recent Oracle database release is 12c; the *c* stands for *cloud computing*, but what is that?

Oracle just does databases, right?

Wrong! At one point in time, if any Oracle product came into an IT shop, it got passed straight to the DBAs for implementation because it had to be database related. Those times are long gone, and understanding the history of Oracle is important so you're not confused by the myriad of Oracle products.

Oracle Corporation started as a database software company but, through remarkable success and a series of acquisitions of other companies, now has a bunch of products. The acquisition of Sun Microsystems yielded SPARC UNIX server hardware, the Solaris operating system, storage systems, and the Java programming language. The purchase of BEA Systems provided the well-respected WebLogic Java application server. Other

purchases, such as PeopleSoft and Siebel, provided Oracle with new applications software.

Today, Oracle products include (but are not limited to) databases, operating systems (Linux and Solaris), database and web application server hardware, disk and tape storage devices, web and application server software, Business Intelligence (BI) software, Java programming language and supporting software, and multiple business applications. As you design your next computer system, you likely will have the option to have most, if not all, the components supplied by Oracle. You can consider the pros and cons of that (end-to-end support versus vendor commitment, for example), based on your mission requirements.

Cloud computing is an architecture where services (such as databases, applications, development environments, or even server infrastructure) are available over the network to the consumer (user). The cloud service provider supplies the product and bills the consumer based on usage. For example, say you wanted to use a type of computer server but you didn't want the overhead of buying and maintaining physical computer hardware and software. A cloud provider could sell you access to that computer server hosted in their cloud. You would pay the cloud service provider a fee based on your usage and the provider would give you access to the servers in their cloud. You get the benefits of having access to the cloud based server without the hassle and costs of buying and maintaining your own servers.

Cloud architecture allows consumers to choose only the services they need and pay for only what they use. Benefits to the consumer include lower cost, not having to provide their own IT service, and generally faster delivery of services than if they built the computing environment themselves.

Oracle 12c is engineered to run on the cloud and be accessible as a service. Through virtualization and multi-tenant architecture, each database application can appear to have its own private database but, in reality, the system is a shared database environment. The use of shared resources and virtualization is a method used by companies and database providers to reduce costs. However, there is no requirement that Oracle 12c must run in the cloud; you may run Oracle 12c just like you run other versions of Oracle without the cloud.

Keeping a Database Safe and Sound

The person who installs and configures a database, manages user access and accounts, ensures space for database objects, and (we hope) makes backups, among many other responsibilities, is the *database administrator*, or DBA.

Medium to large IT shops have one or more DBAs, sometimes even a large team of DBAs. Smaller IT shops may have a person working as a DBA part-time but doing system administration or application development as well. Regardless of whether you're a full-time or part-time DBA, you need to understand the responsibilities and common tasks involved with the care and feeding of a database.



Try not to confuse the DBA with the *application developer*. The process of defining tables, columns, relationships, and rules is application design, data modeling, and ultimately application development. The DBA is peripherally involved with these tasks, but they're typically the application developers' responsibilities.

DBA responsibilities

A DBA has multiple responsibilities around which individual tasks are created to support. Work comes in the form of technical, non-technical, database-centric, and business-centric tasks, but they generally relate to a set of core areas. As a DBA, you have to

- ✓ **Protect the data.** You are the gatekeeper and protector of the data you are entrusted to defend. Data must be protected from theft, damage, and destruction. Protecting the data encompasses both security and backup and recovery. This responsibility commonly comes in the form of establishing, testing, and monitoring backups and implementing security policies and technologies.
- ✓ **Provide access to the data.** You must ensure that the data is available to those users and programs which are authorized to access it. Speedy, quick access is necessary because if a query takes too long to execute and the user gives up, it might as well never finish. Data must be accessible as well; if the database is shutdown or otherwise unavailable, it does the users no good. Providing access to the data involves performance tuning and High Availability (HA).
- ✓ **Maintain the software.** You are responsible for ensuring the general maintenance and housekeeping of the database is kept up to date. Databases must be monitored for errors, ensuring there is enough space to grow as new data is loaded, and monitor processing jobs for success. Furthermore, there is always a new version of software to upgrade to or apply a patch (often after hours). These are generally mundane items

and much of it can be automated via scripts or Enterprise Manager, but they still are your responsibility.

- ✔ **Support the business.** Databases do not exist in isolation. They are part of an overall business objective, and that business objective is usually evolving. New application code is being generated requiring developer assistance, audits of your processes are being conducted, and users always have miscellaneous issues. The most common example is applying a new application release to the database or implementing new technology to enable a new business capability. These efforts periodically correspond to database maintenance, but they are in support of moving the business objectives forward, and the wise DBA supports these efforts.

Common tasks

Life as a DBA varies based on the nature of the business and the IT environment. Here's how the core responsibilities can manifest themselves within a common task:

- ✔ **Protect the data by**
 - Reviewing output logs and messages to ensure the previous night's backups ran successfully.
 - Responding to questions about user access and database auditing procedures. (This task also falls under *support the business*.)
 - Researching and applying Critical Patch Update (CPU) patches to ensure security and vulnerability fixes are made. (This task also falls under *maintain the software*.)
- ✔ **Provide access to the data by** checking Enterprise Manager and your cell phone texts for alerts showing down databases or listeners, which would impact user access to the databases.
- ✔ **Maintain the software by** using Enterprise Manager and log files to search for database alerts, errors, and space usage threshold concerns.
- ✔ **Support the business by**
 - Reviewing output logs and messages to ensure the previous night's application batch jobs and processes completed.
 - Checking e-mail and responding to any application or user issues.
 - Applying application updates to development, test, and production databases to provide additional functionality and bug fixes for users.

Database administration is a respected profession in the IT field. Many people make a comfortable living as an Oracle DBA. Some would say Oracle wouldn't be where it is today if not for the legions of dedicated Oracle professionals worldwide.