

Product Overview



- ◆ DB2 PRODUCTS and PACKAGING
- ◆ DB2 CONNECTIVITY
- ◆ DB2 TOOLS
- ◆ DB2 ADMINISTRATION

*I*n this chapter you will be introduced to the DB2 family of products that run on the Linux, UNIX, and Windows operating systems. This version of DB2 is often referred to as the *distributed* version to differentiate it from the DB2 for z/OS[®] version that runs on an IBM mainframe.

DB2 has the ability to store all kinds of electronic information. This includes traditional relational data, data encoded within XML as well as structured and unstructured binary information, documents and text in many languages, graphics, images, multimedia (audio and video), information specific to operations like engineering drawings, maps, insurance claims forms, numerical control streams, or any type of electronic information. This chapter illustrates some of the ways to access data in a DB2 database using some of the interfaces provided within the DB2 family. A description of each of the DB2 products are provided to illustrate some of DB2's features and functions.

Information as a Service

The DB2 Data Server is an important part of IBM's Information as a Service software portfolio that serves as the atomic level for the broader IBM On Demand architecture.

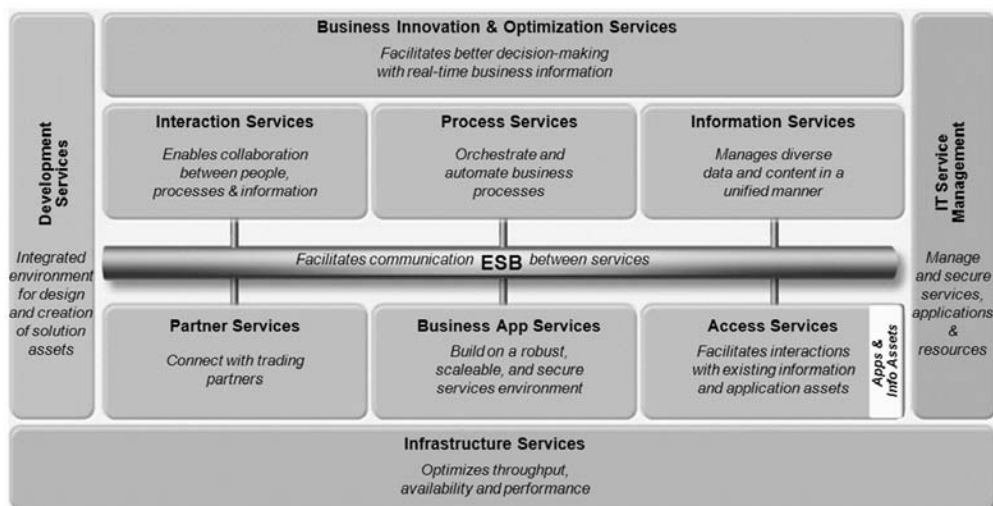


Figure 1–1 *IBM Service Framework for an On Demand business*

In Figure 1–1 you can see that the IBM software portfolio has really evolved into a collection of high value services provided by various IBM software portfolio offerings. The backbone fabric of this IBM reference architecture is the Enterprise Service Bus (ESB) that is used to facilitate communications across this rich set of services.

IT Service Management is mostly provided by various Tivoli® products. The Tivoli portfolio is built around four key disciplines or pillars:

- Security Management
- Storage Management
- Performance and Availability
- Configuration and Operations

Services from these pillars can be used to collectively manage your entire IT framework. For example, Tivoli Storage Resource Manager services can be used enterprise-wide to monitor and report on heterogeneous storage resources to increase storage utilization, identify and resolve potential problems, and ensure application availability through policy-based automation.

Development Services are the culmination of various Rational-based products that are built on the open source Eclipse platform. For example, Rational® ClearCase® provides source control services, and Rational Application Development empowers application developers with a rich set of services that can be used to develop applications, Web pages, and extended custom services for implementation in a Services Oriented Architecture (SOA) or loosely coupled application framework.

Services that enable interaction are typically part of the Lotus® suite of products that enhance collaboration and idea sharing across the enterprise and beyond. Products like Lotus Sametime® Connect can be used for messaging and more.

A number of services in the framework illustrated in Figure 1-1 are provided by the WebSphere® portfolio. For example, a product like WebSphere Integration Developer helps you define business process flows in the standard Business Process Execution Language (BPEL), which are used to implement process services that in turn help you define, orchestrate, and automate business policies. The Enterprise Service Bus (ESB) is provided by the WebSphere ESB product that provides your enterprise services with transformation, transport switching, and routing remediation among other services. Perhaps the most famous product of the WebSphere brand is the WebSphere Application Server that provides a runtime framework for J2EE®-based operations that are part of the Infrastructure Services component.

Finally there's the Information Services which represent the superset of the capabilities you'll learn about in this book. The specific set of services that are typically found in this part of the IBM reference architecture are shown in Figure 1-2.

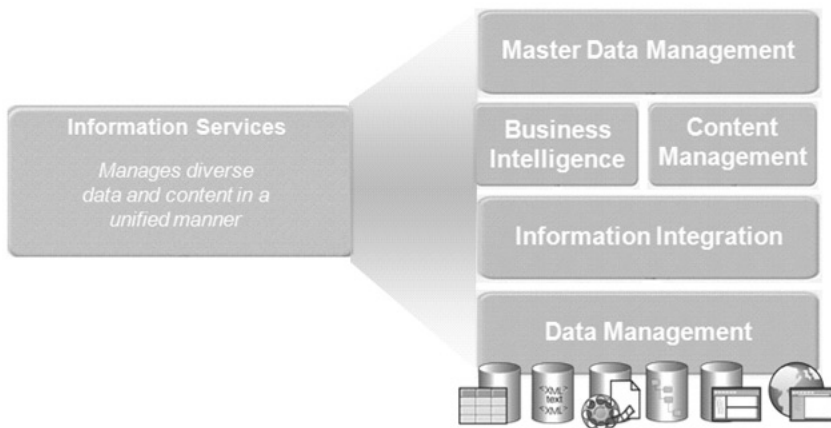


Figure 1-2 *IBM Information Services defined*

The services shown to the right in Figure 1–2 are hierarchical in nature. In other words, as you work from bottom to top, the services provided become richer and more business oriented.

For example, Master Data Management services are provided by the WebSphere Product Center and WebSphere Customer Center products. Master data are facts that describe your core business entities: customers, suppliers, partners, products, bill of materials, employees, and so on. The discipline of Master Data Management seeks to decouple master information from individual applications spread across the enterprise and create a central, application independent, resource. The end result is a simplification of ongoing integration tasks and new application development. This discipline addresses key issues such as data quality and consistency proactively rather than “after the fact”; for example, in a data warehouse (a lower service in this taxonomy). There is also a set of entity resolution services that fit within the Master Data Management service tier.

Business Intelligence services are provided by the DB2 Data Warehouse editions that you’ll learn more about later in this chapter. Content Manager services are provided by the set of Content Management products and are used for document management, archiving, regulatory retention, and are a basis by which unstructured information (such as FAXes, video, voicemail, and so on) can be searched and folded into the information asset.

Information Integration services seek to provide enterprises with ways to share, place, publish, cleanse, and enrich data found in the lower-level data management services. WebSphere Federation Server and its parent WebSphere Information Server are two such products that help implement these services.

Finally, the Data Management services tier is the foundation upon which the other services are built. IBM has a number of data servers that fit into this tier, including DB2, Informix[®], IBM Cloudscape[™], U2, and IMS[™].

This book is specifically about DB2 in this service tier. Specifically, you'll learn how DB2 can provide any number of the high-value data-centric services shown in Figure 1-3.

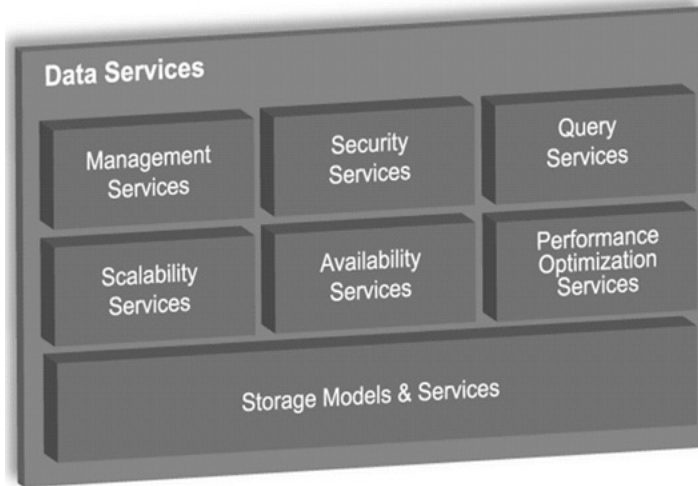


Figure 1-3 *The data services provided by DB2, the focus of this book*

For more information on the entire IBM software portfolio and how they are mapped to the illustrated services shown Figures 1-1 and 1-2, refer to the IBM Web site at www.ibm.com for more details.

The DB2 family of data servers executes on Windows, Linux (which can be run on the entire spectrum of IBM's hardware: System i™, System z™, System x™, and System p™), Solaris™ (both SPARC®-based and Intel®/AMD™-based installations), HP-UX™ (both PA-RISC™-based and Itanium-based installations), i5/OS®, VSE/VM, z/OS, and on pervasive platforms (like Windows Mobile Edition, Blue-Cat® Linux, Symbian®, Palm OS®, J2ME® platforms like the RIM® Blackberry®, and more).

The DB2 code base is optimized for each platform to ensure maximum performance and integration. DB2 for Linux, UNIX, and Windows shares about a 98 percent common code base with platform-specific optimizations at the operating system interface (OSI) level (Figure 1-4).

This means that once you've learned how to administer a DB2 for AIX system, for the most part you'll know how to manage DB2 for Linux or DB2 for Windows; this is the reason why there is a single DB2 certification for all the distributed platforms.

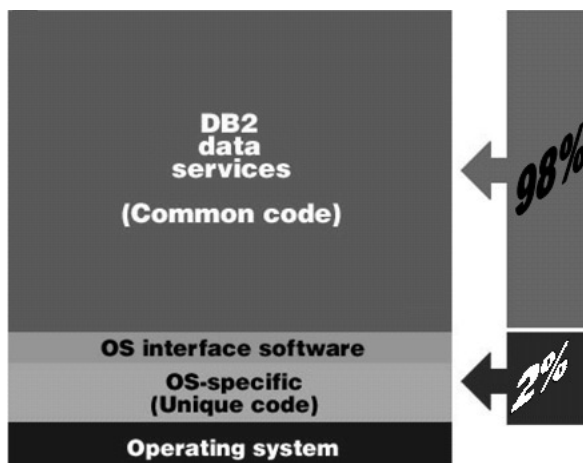


Figure 1–4 *The DB2 code for Linux, UNIX, and Windows is virtually the same*

DB2 for i5/OS and DB2 for z/OS are optimized for their respective environments. For example, DB2 for z/OS is based on a shared-everything disk architecture where the hardware-assisted Coupling Facility is used to serialize access to the shared disk. No such hardware exists for Linux, UNIX, and Windows, and therefore DB2 on these platforms uses a shared-nothing architecture. For this reason, administration tends to vary between these platforms (though many concepts and features are similar). However, the SQL API is 98% common to all the platforms where DB2 runs, allowing applications written on one platform to be easily ported to another. This means that you can build an application on DB2 for Windows and port it effortlessly to DB2 for z/OS. If you build your application according to the *SQL Reference for Cross-Platform Development* handbook, your application will be 100% portable across the DB2 family.

There are other synergies among DB2 running on Linux, UNIX, and Windows, as well as the other DB2 family members. For example, the JDBC driver used for DB2 for z/OS is exactly the same code as is used for DB2 for Linux, UNIX, and Windows. So while there may be variations in specific data definition language (DDL)-based tasks, the data manipulation language (DML) and client APIs are similar.

The DB2 for Linux, UNIX, and Windows Data Server

In the distributed environment, DB2 is available in a number of different packaging options, called *editions*. Furthermore, DB2 is also available as part of other packages that contain additional features and tooling rather than just the base data services provided by DB2.

The mainstream DB2 editions are shown in Figure 1–5:

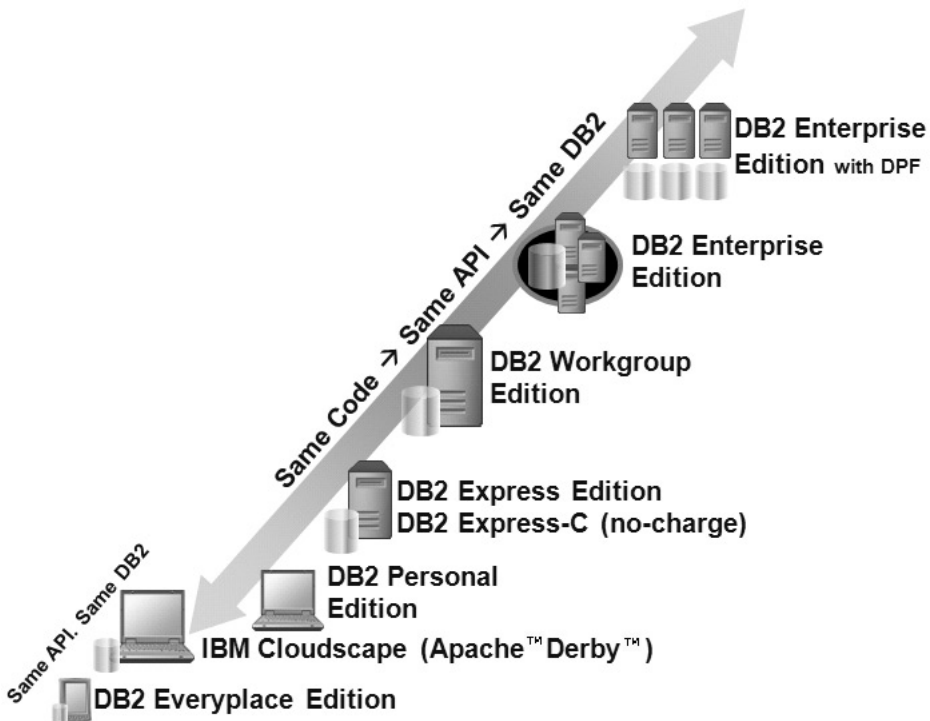


Figure 1–5 *The distributed DB2 family*

For the most part, each edition builds on its child in this hierarchy. For example, if a feature or functionality is available in DB2 Workgroup Edition, it's likely that it's also a part of a higher-level edition, like DB2 Enterprise Edition.

DB2 Note

The packaging and licensing of the DB2 product is very dynamic in nature. This chapter details how DB2 is licensed and packaged as of the time this book was written and is consistent with the focus area of the certification exam. Subsequent point releases could conflict with the information in this chapter, but the exam questions only change with versions.

DB2 Everyplace Edition

DB2 Everyplace (DB2e) is a tiny “fingerprint” database that’s about 350K in size. It is designed for low-cost, low-power, small form-factor devices such as personal digital assistants (PDAs), handheld personal computers (HPCs), and embedded devices. DB2e runs on a wide variety of handheld devices, with support for Palm OS 5.x, Windows Mobile 2003 for Pocket PC, Windows Mobile 2005 for Pocket PC, Windows CE.NET, traditional Windows desktop platforms, Symbian OS Version 7/7s, QNX® Neutrino® 6.2, Linux distributions running with the 2.4 or 2.6 kernel, embedded Linux distributions (like BlueCat) running with the 2.4 or 2.6 kernel, and more.

DB2 Note

DB2e also goes by the name *Mobility on Demand*; if you buy DB2e through an add-on feature pack for DB2 Enterprise Edition (discussed later in this chapter), it’s called Mobility on Demand. If you purchase this product on its own, it’s known as DB2 Everyplace.

The SQL API used to develop DB2e applications is a subset of that used for building full-fledged DB2 data server applications. This means that enterprise applications, for the most part, can be easily extended to include mobile devices. More importantly, it means that if you have DB2 skills, you have DB2e skills. In addition, DB2e is extremely flexible for developers, with support for Open Database Connectivity (ODBC), Java Database Connectivity (JDBC), .NET (including the ADO.NET 2.0 API), and the DB2 Call Level Interface (CLI) APIs.

DB2e is a very simple-to-use data server that requires virtually no maintenance. Typical database administrator (DBA) operations like reorganizations and statistics collection are all performed automatically. Another nice thing about developing DB2e applications is that the database engine is platform independent, so it provides flexibility: You can seamlessly move DB2e databases between devices. For example, you could move a DB2e database populated on a Pocket PC device to a

Symbian smartphone, or whatever other supported device you have, without the need to do anything. This feature, coupled with the rich support for application development, enables developers to quickly build, deploy, and support mobile applications on all platforms.

DB2e is available in two editions: DB2 Everyplace Database Edition (DB2e DE) and DB2 Everyplace Enterprise Edition (DB2e EE). The database component of DB2e DE is the same as DB2e EE; however, DB2e DE has no synchronization middleware to extend or synchronize data to back-end enterprise data servers (although it does come with command line-based import and export utilities). DB2e DE is primarily used for applications that require an embedded database or a local relational storage facility that is exposed to endusers through some sort of application (they never really see the database) yet have stringent footprint requirements because of the device.

DB2e EE distinguishes itself from DB2e DE in that it comes with a data synchronization component called the DB2e Synchronization Server (DB2e Sync Server). The DB2e Sync Server allows you to manage subscriptions and security controls for data that is distributed wirelessly to your hand-held devices and manage data changes on the client devices back into the data center. The DB2e Sync Server also comes with facilities for conflict resolution, application deployment, device identification controls, management policies, and more.

The DB2e Sync Server can synchronize DB2e and Apache Derby/IBM Cloudscape data servers with back-end JDBC-enabled compliant data servers (for example, DB2, Oracle, Informix, SQL Server™, and so on). In addition, there is a special DB2 family synchronization adapter that uses the Data Propagator™ (DPROPR) SQL-based replication technology (which is included in the distributed version of DB2).

The number of concurrent synchronizations that the DB2e Sync Server can support is dependent on the hardware configuration of that server, the associated workload, and data change rates. If you need to scale to handle very large concurrent synchronizations, you can install any Java application server (like IBM WebSphere Application Server). DB2e also supports enhanced scalability and high-availability through its support for DB2e Sync Server farm configurations that allow you to cluster a number of DB2e Sync Servers to provide load balancing and high-availability services.

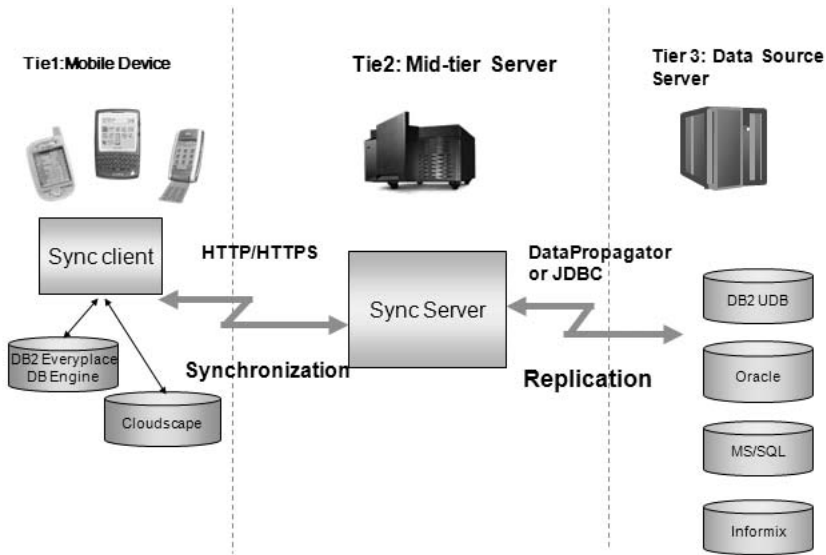


Figure 1–6 A DB2e Enterprise Edition environment

In Figure 1–6 you can see the flow of data in a DB2e EE environment. For example, data is pulled from a database in Tier 3 (the far right of the figure) and placed on a mobile device in Tier 1 (the far left). Tier 1 is typically composed of occasionally connected clients that operate on data and then use the services provided by Tier 2 (the middle of the figure where the DB2e Sync Server resides) to push those changes back to Tier 3. Tier 2 handles issues like conflict remediation and subscription management to ensure that the data quality is maintained throughout its lifecycle until it's at rest.

Apache Derby/IBM Cloudscape

In 2005, IBM donated \$85 million worth of relational database management system (RDBMS) code to the open source community, and the Apache Derby database was born. Apache Derby and IBM Cloudscape are the same databases; the difference is that IBM Cloudscape is sold by IBM with IBM's award-winning 24*7 support and has some add-on features as well.

If you hadn't heard of IBM Cloudscape before the donation news, you'll probably be surprised to learn how many partners, customers, and software packages use this data server. In fact, more than 80 different IBM products use the IBM Cloudscape data server for its portability, easy deployment, open standards-based Java engine, small footprint, and more. IBM Cloudscape is a component that is transparent to

products such as WebSphere Application Server, DB2 Content Manager, WebSphere Portal Server, IBM Director, Lotus Workplace, and many others.

IBM Cloudscape is a Java-based RDBMS that has a 2MB footprint. It's compatible with DB2, supports advanced functions (such as triggers and stored procedures), is easy to deploy, and requires no DBA effort. These same characteristics hold true for the open source Apache Derby as well.

We chose to include the Apache Derby/IBM Cloudscape data servers in this discussion because their SQL API is 100% compatible with the DB2 data server editions in Figure 1–5. This means that you can take any Apache Derby/IBM Cloudscape database and application and move it to a full-fledged DB2 data server if you need more scalability, or you need to take advantage of features that aren't found in these data servers. In fact, a component of DB2 9, called the DB2 Developer Workbench, provides a built-in facility to migrate Apache Derby/IBM Cloudscape schemas and data to a DB2 data server.

DB2 Personal Edition

DB2 Personal Edition (DB2 PE) is a full-function database that enables single users to create databases on their workstations. Since it's limited to single users (it doesn't support inbound client request for code), it's generally not referred to as a data server (although the DB2 engine behind DB2 PE is that same DB2 engine for all editions in Figure 1–5). This product is only available on Linux and Windows. DB2 PE can also be used as a remote client to a DB2 data server. Applications written to execute on DB2 PE are fully portable to the higher-level editions of the DB2 family in Figure 1–5.

DB2 PE is often used by end users requiring access to local and remote DB2 databases, or developers prototyping applications that will be accessing other DB2 databases. In addition, since it includes the pureXML™ technology free of charge, DB2 PE is also a good choice for those looking to acquire DB2 9 pureXML skills. In many cases, because it includes replication features, DB2 PE is used for occasionally connected applications (like field research, sales force automation, and so on) where a richer feature set is required than what's offered by DB2e or Apache Derby/IBM Cloudscape.

The DB2 Express and DB2 Workgroup Editions

Both DB2 Express Edition (DB2 Express) and DB2 Workgroup Edition (DB2 Workgroup) offer the same functions, features, and benefits; they are differentiated only with regard to licensing restrictions, which have a direct affect on the amount of scalability and performance that can be derived from each.

This section will detail these editions, point out differences where they exist, and describe the set of add-on feature packs that you can purchase to extend the scalability, capability, and availability of these data servers.

DB2 Express Edition

DB2 Express is a full-function, Web-enabled client/server RDBMS. DB2 Express is only available for Windows- and Linux-based workstations (unlike DB2 Workgroup). DB2 Express provides a low-cost, entry-level server that is intended primarily for small business and departmental computing. As previously mentioned, it shares the same functions and features as DB2 Workgroup, but is mainly differentiated from DB2 Workgroup by the amount of memory available on the server and the server's Value Unit (VU) rating (which equates to the power of a server's processor cores) on which it can be installed.

DB2 Express can be licensed using the VU methodology, which applies a per-VU charge for the VU rating of a server, or by an Authorized User metric. A DB2 Express server cannot use more than 4 GB of RAM on the server where it is installed. Authorized Users represent individual users that are registered to access the services and data of a single DB2 data server in the environment. For example, if you had a user that needed to access two different DB2 Express 9 data servers and wanted to license this environment with Authorized Users, that single user would require two DB2 Express Authorized User licenses (one for each server).

You can also license DB2 Express using the VU model. No matter what licensing methodology you choose, you cannot install DB2 Express on a server with more than 200 VUs.

DB2 Note

In September 2006, IBM Software Group (IBM SWG) announced a new licensing mechanism called Value Units (VUs) — for those seeking to deploy software for Internet use, or across environments where it isn't possible or feasible to identify users. Previous to this announcement, IBM SWG used to license its software products via a processor license. With the advent of dual core processors, each with different attributes, IBM converted the aggregated processor licensing metric to the more granular VU model. In this model, a core is converted to a number of VUs (referred to as the *VU rating of the server*) for which you are required to purchase the corresponding amount of VUs for the software you want to license. For example, an Intel dual core processor converts to 50 VUs per core. If you had a two-way Intel dual core server, you'd have to buy 200 VUs of DB2. You can learn more about VU licensing at www.ibm.com/software/sw-lotus/services/cwepassport.nsf/wdocs/pvu_table_for_customers.

DB2 Express can play many roles in a business. It is a good fit for small businesses that need a full-fledged relational data server. A small business may not have the scalability requirements of some more mature or important applications, but they like knowing they have an enterprise quality data server backing their application that can easily scale (without a change to the application) if they need it to. As noted, an application written for any edition of DB2 is transparently portable to another edition on any distributed platform

DB2 Express-C: The Little Data Server that Could

DB2 Express-C isn't considered a "real" edition of DB2; however, we chose to include it here because it's very likely the case that this is the copy of DB2 that you're using to learn DB2.

DB2 Express-C is a free data server offering from IBM that you can download from www.ibm.com/software/data/db2/udb/db2express/ or order for free as part of the DB2 9 Discovery Kit. At the heart of the DB2 Express-C is the same scalable and robust data engine that you'll find in the other editions covered throughout this chapter. DB2 Express-C is optimized for two-way dual core servers with no more than 4 GB of memory. Because of the specific optimization in the code base for this architecture, many consider these the "license limitations" for this product.

DB2 Express-C was designed for the partner and development communities, but as you get to know this version, you'll realize it has applicability almost anywhere: as a student trying to learn or get certified in DB2, a hobbyist, and even large enterprises will find this product useful in their environments.

A defining characteristic of DB2 Express-C is that it's generally considered to be a *no limits* data server. DB2 Express-C doesn't have the limits that are typically associated with other competitor's free offerings (ironically, they also carry the Express moniker). For example, there is no database size limit with DB2 Express-C, you can address a 64-bit memory architecture, there are no limits on concurrency or on built-in self-managing features, and more. Where limits do exist, they are more than generous for the workloads for which DB2 Express-C has been optimized to run.

The main features that are *not* included in DB2 Express-C when compared to DB2 Express are:

- Support for high-availability clustering
- The ability to enhance the capabilities of the core data server using add-on feature packs (more on these later).
- Replication Data Capture
- 24x7 IBM Passport Advantage® support model (a special packaging of DB2 Express-C, called DB2 Express-C Fixed Term License, offers this support for

DB2 Express-C servers, but details on this option are outside the scope of this book).

If you want to use any of these features in your environment, you need to, at a minimum, purchase DB2 Express.

For the most part, all of the features found in DB2 Express-C are also available in any of the higher editions found in Figure 1–5. The exception for this product is the pureXML component. To help proliferate and grow XML skills across the database community, IBM generously decided to make this feature available for free with the DB2 Express-C data server. You'll note as you read this chapter that this feature is a chargeable add-on feature pack for all other DB2 data servers.

DB2 Workgroup Edition

DB2 Workgroup is also a full-function, Web-enabled client/server database. Unlike its DB2 Express-C and DB2 Express counterparts, it is available on all supported flavors of UNIX (AIX, HP-UX, and Solaris), Linux, and Windows — this is the main non-resource differentiator between DB2 Workgroup and DB2 Express.

DB2 Workgroup provides a low-cost, entry-level server that is intended primarily for small business and departmental computing. DB2 Workgroup supports all the same features as DB2 Express. Additional features and capabilities can also be added, via feature packs, without having to purchase DB2 Enterprise.

DB2 Workgroup can be licensed using the same options as DB2 Express — it only differs with respect to the architecture limits, which in turn optimize it for specific workloads. For example, the RAM limit for DB2 Workgroup is 16 GB, which is four times the amount you are entitled to use with a DB2 Express data server (which generally translates into better performance, or more supported users). The VU restriction is also more generous. DB2 Workgroup cannot be installed on a server that has a rating of more than 400 VUs, whereas the limit for DB2 Express is 200 VUs. DB2 Workgroup can also be licensed via the Authorized User model and shares the same minimum (five Authorized User licenses) as DB2 Express.

DB2 *Note*

In DB2 8 there were two types of Workgroup Editions in the DB2 lineup: DB2 Workgroup Server Edition (DB2 WSE) and DB2 Workgroup Unlimited Edition (DB2 WSUE). DB2 WSE was licensed by a concurrent or named user license, in addition to a base server license. DB2 WSUE was licensed by a processor metric. In DB2 9, these editions merged into one edition: DB2 Workgroup. The named user and server licenses have been replaced by the simplified Authorized User model and the processor license using the VU metric.

DB2 Workgroup can play many roles in a business. It's a good fit for small- or medium-sized businesses (SMBs) that need a full-fledged relational data server that is scalable and available over a wide area network (WAN) or local area network (LAN). DB2 Workgroup is also useful for enterprise environments that need silo servers for lines of business, or for departments that need the ability to scale in the future. As previously noted, an application written for any edition of DB2 is transparently portable to another edition on any distributed platform.

Add-on Feature Packs for DB2 Express and DB2 Workgroup Editions

DB2 Express and DB2 Workgroup come with the unique flexibility to add enterprise-like services (generally found in DB2 Enterprise) without having to buy a more expensive edition of DB2. Generally the price of these feature packs is such that if you only need one or two specific features, you can save money by purchasing the appropriate feature packs instead of purchasing DB2 Enterprise (as long as you remain within the architectural limitations of the DB2 edition you are implementing). This isn't the case with other competitive data server offerings.

Features Packs for DB2 Express and DB2 Workgroup are licensed in the same manner as the underlying data server. In other words, if you licensed your DB2 Express data server using the VU metric, you have to license any add-on feature packs using the VU metric as well.

The following feature packs are available for DB2 Express and DB2 Workgroup data servers:

- **Workload Management Feature Pack**

Allows you to use the Connection Concentrator and the DB2 Governor, as well as install DB2 Query Patroller on a DB2 Express or DB2 Workgroup data server.

The Connection Concentrator is useful for applications where multiple transient connections perform a limited amount of work in intervals. For example, think about a Web-based application where you browse around and selectively choose items to buy. You may be logged onto the system for a longer period of time while you browse potential items you wish to buy, but you're not making the data server work all the time because you're likely reading the page rather than continually clicking buttons. Concentrating a data server connection improves performance by allowing many more client connections to be processed efficiently, and it also reduces the memory used for each connection. This capability is part of a base DB2 Enterprise installation.

The DB2 Governor, also included by default with DB2 Enterprise, is used to reactively monitor the behavior of applications that run against a DB2 data server. Using the DB2 Governor, you can alter the behavior of applications or

the data server by taking corrective actions in response to thresholds that you define in a configuration file. For example, if an application is using too much CPU, you can set a rule that when this threshold is breached, the application is terminated or given less CPU priority.

DB2 Query Patroller (DB2 QP) is used to proactively manage the workload of a data server, the opposite of the reactive DB2 Governor. With DB2 QP, you can define a set of user and group business policies that are proactively monitored. For example, if a user submitted a query that the optimizer estimated would cost 1,000,000 timerons, and you set a business rule stating that no queries can be larger than 100,000 timerons, DB2 QP would stop this query from being processed on the data server. You can also use DB2 QP to perform charge-back accounting because it tracks valuable usage information (this information can also be used for performance tuning and more). DB2 QP is detailed in the “DB2 Query Patroller” section later in this chapter.

- **Performance Optimization Feature Pack**

Makes available the use of materialized query tables (MQTs), multi-dimensional clustering (MDC) tables, and query parallelism for DB2 Express and DB2 Workgroup servers. All of these features are used to provide exceptional performance and are part of a base DB2 Enterprise installation.

DB2 comes with a number of high-performance objects and capabilities that allow it to scale to hundreds of thousands of users and into the millions of transactions per minute or queries per hour. This feature pack provides the ability to create MDC tables and MQTs in your DB2 Express and DB2 Workgroup data servers. These objects provide immense benefits for applications running on DB2. In fact, we’d say that some of the most important components for any high-performing application are part of this feature pack. (If you’re running a data warehouse be sure that you know what MDCs and MQTs are.) If you’re looking to really boost the performance of an application running on the smaller servers for which DB2 Express and DB2 Workgroup were made, this feature pack has components that could prove very valuable to your business.

- **High-Availability Feature Pack**

Gives DBAs a free two-node license of Tivoli System Automation (TSA) for high-availability failover clustering, the ability to run online table reorganizations, and the High-Availability Disaster Recovery (HADR) feature. All of the features in this feature pack are part of a base DB2 Enterprise installation.

HADR is a high-availability feature that provides a database availability protection plan that is very simple to set up and use. The best part about HADR is that you set it up with mere clicks of a button. The online table reorganization capability, as its name implies, allows you to reorganize tables online. Finally, this

feature pack includes a two-node cluster license for Tivoli System Automation (TSA) for AIX and Linux — you can use it to cluster together your servers for high-availability or to automate the failover of an HADR environment.

- **pureXML Feature Pack**

Provides the ability to create pureXML columns in a DB2 Express or DB2 Workgroup data server and use an associated set of XML services when working with this data.

You might be confused about the DB2 XML Extender (covered later in the “DB2 Extenders” section) and the pureXML add-on feature pack that’s available in DB2 9. The DB2 XML Extender provides the XML capabilities that were part of the DB2 8 release. In contrast, the pureXML feature enables DB2 servers to exploit the new hybrid storage engine that stores XML naturally in DB2 9. The performance, usability, flexibility, and overall XML experience of pureXML can’t even be compared to the older DB2 XML Extender technology; however, the DB2 XML Extender is still shipped in DB2 9 free-of-charge. If you are planning to use XML in your data environment we strongly recommended you use the pureXML feature.

The pureXML feature lets you store XML in a parsed tree representation on disk, without having to store the XML in a large object or shred it to relational columns as you are forced to with the DB2 XML Extender. This can be very beneficial for applications that need to persist XML data. Access to XML data via the pureXML feature pack is a very natural experience; for example, you can use SQL or XQuery to get to relational or XML data.

The pureXML feature also has facilities to store XML Schema Definition (XSD) documents in a native XML Schema Repository (XSR) service. It also supports schema annotations for document shredding, validation services, and more.

DB2 *Note*

DB2 9 supports the shredding of XML data to relational in the same manner as the DB2 XML Extender, but it uses a different and far superior technology to do it. You may want to shred your XML to relational for any number of reasons, such as when the XML data is naturally tabular. To shred XML to relational using the DB2 XML Extender, you have to hand-generate Document Access Definition documents that map nodes to columns, and so on. With DB2 9, even without the pureXML feature, you can use the DB2 Developer Workbench (covered later in this chapter) to shred your data and automate the discovery of these mappings. The new mechanism in DB2 9 is also significantly faster than the DB2 XML Extender method.

- Homogeneous Federation Feature Pack

Provides the ability to create nicknames across the DB2 family of data servers. This feature allows developers to build applications that access DB2 tables that reside on different platforms without regard to their location. For example, you could use this feature to easily create an application that performs a join of data that resides on a DB2 for i5/OS data server with one that's running on DB2 for Windows. Even if you were working within an integrated development environment (IDE) such as IBM Rational Application Developer or Microsoft Visual Studio 2005, you still wouldn't be able to tell where each table actually resides — which is the whole point. The capability of this feature pack is a subset of the WebSphere Federated Server product covered later in this chapter.

DB2 Enterprise Edition

DB2 Enterprise Edition (DB2 Enterprise) is the premier data server offering from IBM that is the foundation of many mission-critical systems and the primary focus of this book. It is fully Web-enabled, scalable from single core servers to symmetric multicore servers and to massively parallel systems.

This edition of DB2 is considered to be the full-function offering from IBM (for example, it's the only DB2 edition that includes table partitioning). It is available on the same supported flavors of Linux, UNIX, and Windows as DB2 Workgroup. Applications built for DB2 Enterprise can scale upward and execute on massively parallel systems or downward to smaller servers.

DB2 Enterprise is meant for large and mid-sized departmental servers. DB2 Enterprise includes all the functions of the DB2 Express and DB2 Workgroup editions, and more. Additionally, there is a set of feature packs that are exclusive to this edition, such as the new DB2 9 Storage Optimization feature that provides deep row compression, and more.

DB2 *Note*

Some feature packs, such as the Database Partitioning and the Storage Optimization features, are not available using the authorized user metric.

DB2 Enterprise can be licensed using the same VU metric as DB2 Express and DB2 Workgroup; however, in the case of DB2 Enterprise, there is no RAM or VU server rating limits for the server on which this product is installed. DB2 Enterprise can also be licensed via the Authorized User metric. In contrast to DB2 Express and DB2 Workgroup, when licensing DB2 Enterprise using the Authorized User metric, the minimum number of Authorized Users for which it must be licensed is 25 Authorized Users for every 100 VUs on the server where it is installed. For

example, if you installed DB2 Enterprise on an 8-core System p server (rated at 800 VUs), you would have to buy 200 Authorized User licenses at a minimum to use this licensing option, even if you intended for this data server to support fewer than 200 users.

Add-on Feature Packs for DB2 Enterprise Edition

DB2 Enterprise also comes with a set of add-on feature packs just like DB2 Express and DB2 Workgroup. For the most part (the exception being the pureXML feature pack), the extensibility features that can be purchased for DB2 Enterprise are unique to this edition of the DB2 family and provide even richer enterprise-capabilities to the data run-time environment.

The way you license DB2 Enterprise feature packs must be identical to how the DB2 Enterprise data server was licensed. This is one area where DBAs must pay particular attention when deciding how to license their DB2 Enterprise data servers because some feature packs aren't available with the DB2 Enterprise Authorized User license.

The feature packs that are available in DB2 Enterprise via Authorized User or Value Unit licensing are:

- Performance Optimization Feature Pack

DBAs get licenses to install DB2 Query Patroller and the DB2 Performance Expert tool with this feature pack. The DB2 Query Patroller component of this feature pack is the same one that's part of the Workload Management feature pack available for DB2 Express and DB2 Workgroup.

The DB2 Performance Expert is a separately purchasable tool that's used to simplify performance management and tuning. It offers DBAs a consistent view into their instances, subsystems, databases, and applications across the DB2 family. It has a set of pre-canned reports to identify resource shortage and exception conditions in DB2 including locking conflicts, deadlocks, and application and SQL statements causing high workload. The DB2 Performance Expert also includes a set of detailed reports about SQL, database, and buffer pool activity with trend analysis and what-if hypothesis testing for performance evaluation.

DB2 Note

The version of the DB2 Performance Expert that's part of this feature pack can only be used for DB2 data servers running on Linux, UNIX, and Windows. The DB2 Performance Expert tool when purchased outside of this feature pack can be used across the DB2 family.

- pureXML Feature Pack

This feature pack provides the same capabilities as the one offered for DB2 Express and DB2 workgroup detailed earlier in this section.

- Advanced Access Control Feature Pack

Provides label-based access control (LBAC) protection services to data stored in a DB2 Enterprise data server. Using this feature, data stewards can control the read and write access of a user at the table column and row level. LBAC implements data access controls by creating an affinity between columns and generated protection security labels. If users attempt to get data from a table, they must have matching label credentials (or a parent label) granted to them. When there's a match, access is permitted; without a match, access is denied.

You can use this feature pack to create a security framework whose architecture is built on a hierarchical representation that matches the data access hierarchy, an array of that business entity, or a mix of the two. LBAC can also be used to restrict access to XML documents in columns (though as of the time of this writing, you cannot attach labels to fragments within the document itself).

- Geodetic Data Management Feature Pack

This feature pack includes the DB2 Geodetic Data Management Extender that can be used to provide advanced spatial analysis capabilities. What separates this feature from the free DB2 Spatial Extender that comes with all DB2 data servers is that the Geodetic Data Management feature pack contains a built-in set of algorithms that take into consideration the curvature of objects, such as the earth's surface and so on. For example, maps generally are associated with some sort of applied projection in consideration of the map's purpose. The Mercator projection is very popular for navigational maps. When you look at a map generated using this projection, you'll notice that the top and bottom of the map seem much bigger than they really are. The fact that Greenland is one fourteenth the size of Africa often comes as a surprise when people think back to their public school atlas — this feature pack compensates for these distortions.

While distortions caused by projections may not be of significance to applications that attempt to locate an address or division of a city, for weather pattern analysis or defense programs it could be very significant. If you can't afford to lose accuracy because of a projection, the Geodetic Data Management feature pack may be appropriate for your applications.

- Real-Time Insight Feature Pack

Useful for managing large volumes of incoming data streams. Existing infrastructures can be easily overwhelmed when trying to manage large volumes of incoming data. Incoming data with message rates of tens to hundreds of thou-

sands of messages per second can make it difficult to analyze this high volume of data.

The DB2 Real-Time Insight feature pack is powered by the DB2 Data Stream Engine (not discussed in this chapter since it's beyond the scope of this book) that enables organizations to store and forward high volumes of data from multiple data streams. The data messages from the feed can be aggregated, filtered, and enriched in real time before being stored or forwarded.

DB2 Data Stream Engine can load high volumes of data into the DB2 data server and make that data available to queries in real time through SQL. One example is a financial market data stream that provides information about financial transactions, such as stock trades and quotes.

The benefits of the Real-Time Insight Feature include:

- Scalable solution loads large volumes of data with high throughput and low latency
 - Simultaneous storing and publishing of data from multiple feeds
 - Insight into the data with filtering and aggregation from feeds before storing and publishing
 - Maintenance of metadata, such as current state, for entities that are processed from the feeds
 - Simultaneous persistence of data to multiple database servers on multiple hosts
 - Real-time access by use of shared memory storage
 - Easy access to both real-time and historical data through standard SQL, C-API, and Java API interfaces.
- Mobility on Demand Feature Pack

This feature pack provides the components of DB2 Everyplace Enterprise Edition detailed earlier in this chapter. When using DB2e via the feature pack, you need to be aware that the DB2 SyncServer must be collocated with the DB2 Enterprise data server (this is the reason for its reduced cost when compared to a full licensed version of DB2 Everyplace Enterprise Edition).

- Homogeneous Federation Feature Pack

This feature pack provides the same capabilities as the one offered for DB2 Express and DB2 Workgroup detailed earlier in this section.

Two feature packs available in DB2 Enterprise only through Value Unit licensing (in other words, you can't buy these feature packs with DB2 Enterprise servers that are licensed with authorized users) are:

1. Database Partitioning Feature (DPF)

DB2 Enterprise provides the capability to enable DB2 to partition data across clusters or massively parallel servers. To the end user or application developer, a partitioned database appears to be on a single system, yet SQL statements are processed in parallel across all servers, thus increasing the execution speed for any given query.

The DPF delivers the true principals of scalability to a DB2 Enterprise environment, namely:

- **Double the resources, double the data:** Each partition processes the same amount of data as before, and response times and throughput remain constant.
- **Double the resources, keep data constant:** Each partition processes half the amount of data as before, and response times will be cut in half, and throughput will double.
- **Keep resources constant, double the data:** Each partition processes double the amount of data as before, response times should double, and throughput will be cut in half.

You can partition data using the DPF across logical (within a larger SMP) and physical servers. An example of a partitioned database across multiple physical servers (though each server typically is small SMP server) is shown in Figure 1–7.

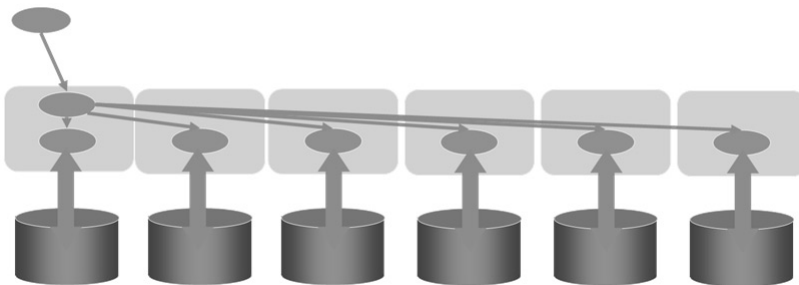


Figure 1–7 DB2 partitioned across multiple servers using DPF

In Figure 1–7 you can see that there are actually six copies of DB2. However, they all appear as a single copy to applications and administrators. Imagine the performance difference between a single copy of DB2 scanning 600,000 rows versus each copy of DB2 owning its own data and resources scanning just 100,000 rows. This is the power of the DPF.

Parallelism in DB2 is automatic and extended to the hash partitioning algorithm used in the DPF. If you are selecting just a few records (where partitioning key value = X), then DB2 will send that query directly to the node that contains that data. However, if you are scanning large amounts of data (as shown in Figure 1-7, typical in data warehousing), then DB2 will send the query to all partitions in the cluster and automatically parallelize the data access operations, driving more resources (RAM, CPU, and I/O) to get the job done faster.

It's not just query performance and faster maintenance operations that are delivered by the DPF. There is a significant resource savings per server because each server owns 1/nth of that data and generally requires fewer resources. Compare the partitioned database (shown on the left side in Figure 1-8) and the non-partitioned database on the right.

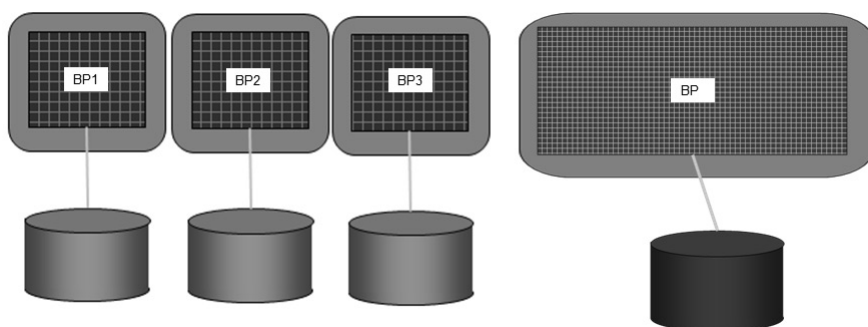


Figure 1-8 *Saving resources with the Database Partitioning Feature*

You can see that the servers comprising the partitioned database on the left side of Figure 1-8 require less memory, as each server is only responsible for 1/nth of the data. In contrast, the non-partitioned database on the right side of Figure 1-8 requires that much more memory be allocated to the buffer pool to accommodate the data.

Finally, DB2 can mix its intra-partition parallelism capabilities (where it runs components of an SQL statement in parallel on a single server) with its inter-partition parallelism (the DPF feature) as shown in Figure 1–9.

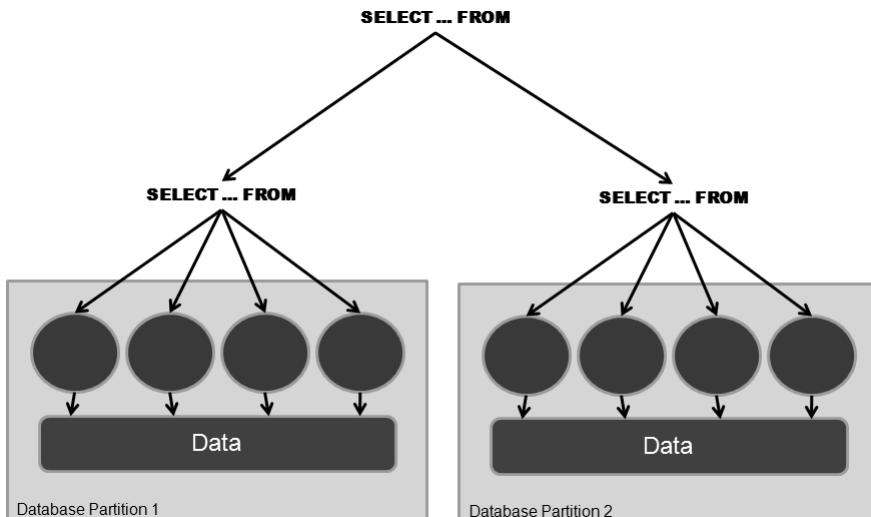


Figure 1–9 *Two kinds of parallelism with DB2 and DPF*

You should consider the DPF feature pack as an add-on to DB2 Enterprise in the following circumstances:

- The speed of database utility operations is key to your business operations. Operations like reorganizations, backups, and so on are parallelized with DPF so they can be performed much quicker. For example, if the six servers in Figure 1-7 all run the backup utility, they each only need to back-up one-sixth of the total data, and this operation should complete in one-sixth of the time when compared to a non-partitioned database.
- You need to shrink your batch window because of long extract, transform, and load processes (ETL). For example, data server load jobs are also parallelized. For example, if you had to load 600 GB of data, only 100 GB of data would need to be loaded into each server.
- Rolling window data update requirements for the warehouse make parallel SQL processing and additional log space essential. Not only does DPF give you more power to process the SQL, it gives you more resources (log space, memory, and so on).
- DPF should be considered if the database contains more than 400 GB of RAW data. Other indicators for using DPF are the total number of rows in the table and whether scan performance is critical to workload performance.

Large databases can be supported with DB2 Enterprise Edition, but large databases typically benefit from the DPF.

- Your environment is characterized by complex queries that involve large aggregations, multi-table joins, and so on. Having multiple servers working on the same SQL problem can generally return results faster than a single server.
- Your servers have plenty of available memory. Even though DB2 has 64-bit memory support for non-partitioned databases, multiple partitions have proven to provide more linear scalability and more efficient usage of memory than SMP parallelism alone.

2. DB2 Storage Optimization Feature

This feature provides deep row compression for your DB2 tables, as well as a backup compression utility that can significantly reduce the on disk size of your database backups. Tests have shown that more than 70 percent of table compression (data only) can be achieved when using this feature pack. Note that indexes remain uncompressed for performance reasons.

A number of other benefits, in addition to disk savings from having smaller tables, arise from row compression. Backups will be smaller (even without backup compression); maintenance operations (like backup) should run faster since there are fewer data pages to backup up; Q/A and test environments will be smaller; heating, ventilation, and air conditioning (HVAC) charges will be decreased since you need fewer disks; and query performance likely will increase because each I/O to disk brings back more rows into memory.

Row compression will drive up the CPU utilization of a data server. However, most systems (especially data warehouse systems) are I/O bound and, since compression will allow more rows on a data page and therefore more rows in the resident memory buffers, you'll likely see improved overall performance of your application.

Backup compression is also a part of this feature pack. The database backup architecture in DB2 is based on an open pluggable interface that allows you to use either the default compression algorithm that's shipped with DB2 or one of your own. The compression algorithm you use is embedded within the backup image so that it can be restored on a different server. Quite simply, if disaster happens and you lose the server where the compression algorithm is located, you can still get your data back because the compression algorithm is part of the backup image.

DB2 Connectivity

DB2 provides a variety of options for connecting to other DB2 and non-DB2 data servers:

- **DB2 Clients** — DB2 client code is required on workstations for remote users to access a DB2 database.
- **DB2 Drivers** — Some APIs that are part of a DB2 client are packaged on their own, outside of a client installation. DB2 drivers offer a more lightweight deployment option for specific API connectivity.
- **DB2 Connect™** — This add-on product provides support for applications executing on Linux, UNIX, and Windows to transparently access DB2 data servers on i5/OS, VM/VSE, and z/OS environments. DB2 Connect provides much more than just connectivity; more details of its capabilities are provided later in this chapter.

DB2 Note

DB2 Connect is not required to access DB2 for any of the LUW platforms.

- **DB2 Replication** — This feature provides replication capabilities for DB2 data servers. There are two kinds of replication, SQL-based replication and Queue-based replication (also known as Q-replication). Both are sometimes referred to as Data Propagator (DPROPR).

SQL-based replication is included in all of the DB2 mainstream editions that run on Linux, UNIX, and Windows. It's made up of two components, CAPTURE and APPLY. It uses SQL to replay data on target servers.

Q-Replication is built on the IBM WebSphere MQSeries® technology and is generally thought to be a more available and powerful replication technology. You can add this capability to a DB2 environment through one of the WebSphere Information Integrator products.

- **WebSphere Information Server or WebSphere Federation Server** — The WebSphere brand has a suite of products that help you publish, place, cleanse, transform, enrich, and access data across heterogeneous data sources.
- **WebSphere Application Server** — A part of this application server is shipped with DB2. It allows developers to use Java as a platform in a transaction processing environment.

DB2 Clients

Once a DB2 application has been developed, you need to install connectivity software on each client workstation in a two-tier environment. If you are deploying a Web-based application in a three-tier environment, you need to, at a minimum, install the DB2 client connectivity software on the application server. In DB2 9 there are basically two types of clients, the *DB2 Runtime Client* and the *DB2 Client*.

A DB2 Runtime Client provides the minimum *client* footprint (about 20–30 MB) to support connectivity to DB2 9 data servers (the exception to this is if you choose to support communication via specific API drivers, covered later in this section). When you install a DB2 Runtime Client, you install a driver to support all the common programming methodologies, including ADO.NET, ADO, OLE DB, ODBC, JDBC, SQLJ, static SQL, and more.

DB2 Note

In this chapter, the use of the term *client* with a lowercase c refers the DB2 client software in general, while an uppercase *C* refers to a specific DB2 client package.

Using a DB2 Runtime Client you can perform basic functions, such as running any DB2 commands or SQL statements from a DB2 CLP or CLP session; however, for the most part, this client comes with no other capabilities than supporting data connectivity. For example, there are no graphical-based management tools, documentation, and so on. You can download a DB2 Runtime Client at www14.software.ibm.com/webapp/iwm/web/preLogin.do?lang=en_US&source=swg-dm-db2rtcl.

DB2 Note

The DB2 8 Runtime Client is very different from the DB2 9 Runtime Client. In DB2 8, there was a special client only available on Windows called the *DB2 Runtime Client Lite*. This client had a reduced footprint compared to the DB2 9 Runtime Client and didn't include tooling, documentation, and so on. The DB2 9 Runtime Client is analogous to the DB2 8 Runtime Client Lite, only is now available on all supported platforms.

A special DB2 Runtime Client is made available as Windows Merge Modules (.msi files), which makes the installation of this software within Windows applications more natural. You can download these files at <http://www-304.ibm.com/jct03002c/software/data/db2/runtime.html>.

In contrast, the DB2 Client includes all the functions found in the DB2 Runtime Client plus functions for client-server configuration, tools for database administration and application development, samples, and more. For example, this client includes the Configuration Assistant that provides graphical administration and connectivity tools, as well as a host of programming samples, and so on. The DB2 9 Client replaces the functions found in both the DB2 8 Application Development and DB2 8 Administration clients.

A DB2 Client's footprint is directly correlated to the components you select to install and can take between 200 MB and 800 MB depending on the options you select. The process to install a DB2 Client is very similar to a DB2 data server. You can download a DB2 Client at <http://www-304.ibm.com/jct03002c/software/data/db2/runtime.html>.

The choice of which DB2 client to install should be based on the requirements of the application on the client machine. For example, if you have a database application developed for Linux, UNIX, or Windows and you do not require the DB2 administration or application development tools from a Windows workstation, yet want to support applications written in multiple programming languages like .NET and Java, you should install the DB2 Runtime Client on that workstation.

Some enterprises prefer to deploy the DB2 Runtime Client code remotely on a dedicated server. In these environments, remote workstations need to access the DB2 Runtime Client code remotely from a code server before accessing to DB2. This type of configuration is known as a thin client. A thin-client configuration can provide many benefits, such as a central catalog directory of all database connections and a single footprint of code to manage. The trade-off is that clients must load the client *.dlls* from the code server before accessing the remote data server. It should be noted however that this performance hit is only "noticed" on the initial connection. Once the client code is loaded on the client workstation, subsequent calls to the code server are not needed.

When you want to deploy your DB2 application, you only need to ensure that a DB2 Runtime Client is installed on each workstation executing the application. Figure 1-10 shows the relationship between an application, a DB2 Runtime Client, and the DB2 data server. If the application and database are installed on the same server, connectivity is considered to come from a local client. If the application is installed on a system other than the DB2 data server, its connectivity is considered to come from a *remote client*.

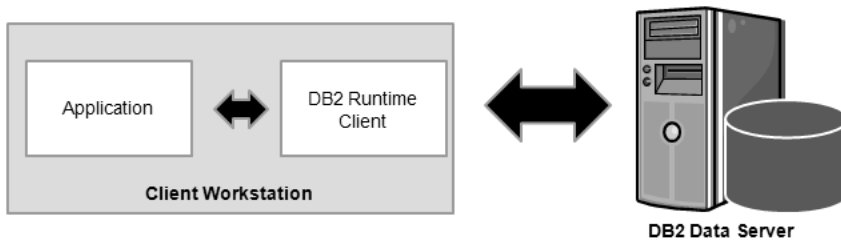


Figure 1–10 *Accessing a DB2 data server using the DB2 Runtime Client*

DB2 client connectivity can be configured using various supported communication protocols. The supported protocols vary according to operating system:

- TCP/IP — used in all environments
- Named Pipe — used in Windows environments

DB2 Drivers

Some larger enterprises and independent software vendors (ISVs) want to deploy their applications written in single language without the overhead of installing and maintaining a DB2 client on each client workstation. Despite the much smaller footprint of the DB2 Runtime Client in DB2 9, IBM makes available two drivers that can be deployed for connectivity on their own, outside of a DB2 Runtime Client or DB2 Client installation.

The *IBM Driver for JDBC and SQLJ* can be deployed to support Java-based connectivity to a DB2 data server. This driver is about 2 MB in size and can easily be embedded within your application. In fact, it comes with a royalty-free distribution license for this very purpose.

You should note however that although this driver can connect to the entire DB2 family, if you're connecting to a DB2 for i5/OS, DB2 for VM/VSE, or DB2 for z/OS data server, you need to additionally license this connectivity with DB2 Connect (covered in the next section) to ensure you are compliant with respect to licensing.

It's important to note that the IBM Driver for JDBC and SQLJ solely supports Java applications. Aside from not including tooling and documentation, this driver doesn't support data server connectivity using other APIs. For example, you can't support a .NET application using this driver. You can download this driver at www14.software.ibm.com/webapp/download/preconfig.jsp?id=2004-09-20+10%3A09%3A21.003415R&cat=database&fam=&s=c&S_TACT=105AGX11&S_CMP=DB2.

DB2 9 includes a new standalone driver for CLI and ODBC applications called the *IBM DB2 Driver for ODBC and CLI*. This driver delivers the same benefits and restrictions as the IBM Driver for JDBC and SQLJ to your DB2 environment, except it solely supports CLI/ODBC connections. You can download this driver at www14.software.ibm.com/webapp/download/preconfig.jsp?id=2004-09-20+13%3A44%3A48.813589R&cat=database&fam=&s=c&S_TACT=105AGX11&S_CMP=SPLT.

The main benefits of the IBM DB2 Driver for ODBC and CLI are:

- Light-weight deployment solution for ODBC/CLI applications
- Much smaller footprint (25M installed) than DB2 RTCL and DB2 Client
- No need to have DB2 RTCL or DB2 Client installed on the client machines
- Same concept as the JDBC/SQLJ driver but for ODBC/CLI applications

DB2 Connect

DB2 Connect is a separate product family that licenses client connectivity from distributed platforms to the DB2 family running on i5/OS, VM/VSE, and z/OS operating systems (hereafter referred to as *mainframe*). Although communications between any members of the DB2 family use the Distributed Relational Database Architecture (DRDA), you have to explicitly license access from DB2 for Linux, UNIX, and Windows clients to these data servers.

A client accessing DB2 running on the mainframe is generally referred to as a *DRDA Application Requester (DRDA AR)* and the DB2 server that manages the client connection as a *DRDA Application Server (DRDA AS)*. DB2 Connect only uses TCP/IP as the transport protocol between a DRDA AR and DRDA AS.

DB2 Note

If you are connecting to DB2 for Linux, UNIX, and Windows from a client running on a mainframe, you don't need DB2 Connect. DRDA AS functionality is built into DB2 data servers running on distributed platforms.

Some of the major capabilities provided by DB2 Connect include:

- Support for programming APIs such as ADO, ODBC, OLE DB, CLI, JDBC, SQLJ, ADO.NET (and more) to DB2 running on the mainframe.
- Federation across the DB2 and Informix data servers. You can extend the reach of this federation capability to heterogeneous sources by adding either WebSphere Federation Server or WebSphere Information Server.
- Connection pooling and the more powerful connection concentration services

that provide minimal resource consumption on the mainframe.

- System z Sysplex exploitation for failover and load balancing. For example, DB2 Connect includes high-availability features such as z/OS Workload Manager Integration (WLM) and automatic re-direct to a data-sharing group.
- Integration into mainstream development environments for mainframe data access. For example, DB2 Connect makes data artifacts on the mainframe (including nonrelational-related services like CICS and VSAM) transparent to the world's most popular IDEs; for example, a .NET developer can use Microsoft Visual Studio 2005 and build stored procedures without knowing where the data actually resides.

DB2 Connect is often used to enable mainframe resources for Web access, better integrate legacy and heritage systems with new deployments on distributed platforms, modernize the application development experience for mainframe development (for example, moving from COBOL to Java), and off-loading development cycles to lower-cost operational environments.

Depending on the DB2 Connect edition you purchase, you can create a connectivity architecture for single- or multi-tier environments.

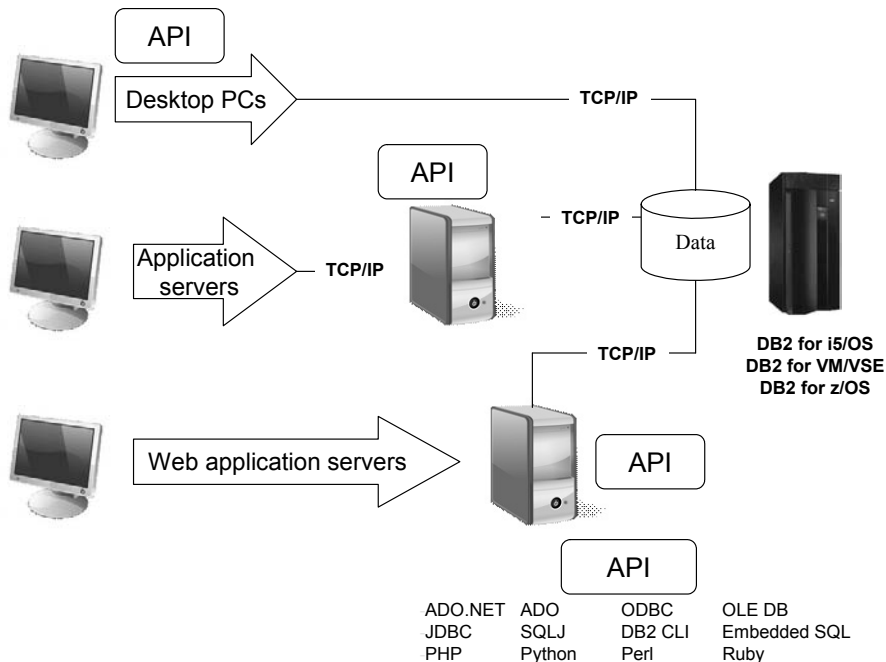


Figure 1–11 *Using DB2 Connect is a single-tier or multi-tier environment*

You can see in Figure 1–11 that there are many ways to implement a DB2 Connect product. In a multi-tier environment, you use DB2 client software to connect to the DB2 Connect gateway (in a client/server model) passing the connection context to the mainframe resource. In a Web-based deployment, you may have DB2 Connect on its own or collocated with the application server. Depending on your configuration, you may or may not need to install a DB2 client.

DB2 Connect products can be added on to an existing DB2 data server installation, or act as a standalone gateway. Either way, it's purchased separately from DB2. If you are using one of drivers discussed earlier in this chapter, technically you can make the connection to a mainframe DB2 resource, but you still need to purchase a valid DB2 Connect license. There are a number of DB2 Connect editions available and each is suited for a particular implementation of the DB2 Connect software. With the exception of DB2 Connect PE, all of the DB2 Connect editions offer the same function and features; they are merely differentiated by capability. All editions of DB2 Connect PE are considered *gateways* since they provide client/server access to workstations. To connect to a DB2 Connect gateway, you simply use a DB2 client.

DB2 Connect Personal Edition (DB2 Connect PE) is similar to DB2 Personal Edition in that it is used for single workstations and doesn't provide its users with any server capabilities. Using DB2 Connect PE, you can make direct connections to DB2 running on the mainframe. In Figure 1–11, the desktop users likely have DB2 Connect PE installed in their workstations. DB2 Connect PE can also act as a regular DB2 client and connect to distributed versions of the data server as well. Some environments deploy DB2 Connect with both direct and gateway connectivity. For example, a developer may make direct connections using the DB2 Connect PE software for certain stages of the development cycle. But once the application is in production, they may use the DB2 Connect PE software to connect to a DB2 Connect gateway to take advantage of the features it provides, such as load balancing, connection optimization, high-availability, and more. DB2 Connect PE is licensed on a per-workstation basis.

DB2 Connect Enterprise Edition (DB2 Connect EE) can be licensed using two different user-based options: the number of Authorized Users or the number of Concurrent Users. DB2 Connect EE provides the ability to implement a three-tier connectivity architecture that allows you to isolate connectivity to mainframe resources to a specific tier in your architecture. Again, DB2 clients (or drivers) are used to connect to the gateways, and then the DB2 Connect software manages the connection context to the mainframe.

DB2 Connect Application Server Edition (DB2 Connect ASE) is licensed by the total VU rating of all the applications servers that connect to the mainframe. The VUs that you purchase have nothing to do with how many DB2 Connect servers

you set up, or what their rating is. For example, if your Web server farm was rated at 1,000 VUs and you set up a DB2 Connect gateway farm rated at 5,000 VUs, you would have to buy 1,000 VUs of DB2 Connect ASE. DB2 Connect ASE is well suited for environments where you expect to see more growth of mainframe resource than your Web tier because its costs are directly linked to the VU rating of the Web serving tier.

DB2 Connect Unlimited Edition (DB2 Connect UE) is available for i5/OS and z/OS. DB2 Connect UE for i5/OS is licensed by the number of managed processors attached to the i5/OS partition while DB2 Connect UE for z/OS is licensed with a base server license plus a per-unit charge for the MSU rating of the System z server you are connecting to. DB2 Connect UE is well suited for environments that have flat mainframe cycle growth and heavy Web tier growth as the costs of DB2 Connect UE are directly correlated with the capacity of the mainframe resource. DB2 Connect UE comes with a free copy of Mobility on Demand to move mainframe data to occasionally connected devices, as well as DB2 Connect PE. Essentially, this edition of DB2 Connect, as its name would imply, allows for unlimited deployments.

There's so much more to the DB2 Connect product than what's detailed in this section. In fact, the name DB2 Connect is misleading because connectivity is just a small part of what this product can do; however, these details are outside the scope of this book.

DB2 Replication

DB2 Replication allows for data to be propagated from one location to another. SQL-based replication, also known as Data Propagator (DPROPR) replication, is a free component of the DB2 for Linux, UNIX, and Windows data servers (it's a paid feature on the mainframe). You can also extend this replication capability to support a wide variety of data servers, including DB2, Oracle, Microsoft, Sybase, Informix, IMS, Lotus Notes, and flat files with WebSphere Replication Server.

DB2 Note

SQL replication is also the core technology behind the DB2e SyncServer that enables mobile users to keep their data synchronized with corporate data.

An example of SQL-based replication is shown in Figure 1–12. SQL-based replication has been around for over a decade and has proven itself to be extremely flexible, resilient, easy to set up, and scalable.

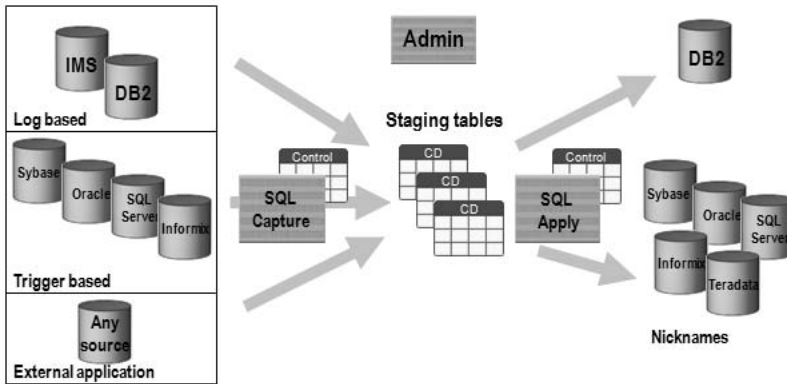


Figure 1-12 *SQL-based replication extended with WebSphere Replication Server*

In Figure 1-12 you can see the broad reach of the SQL-based replication technology when it is extended with the WebSphere Replication Server product. As you may recall, distributed versions of DB2 come with the DPROPR technology built in and can support replication subscriptions across the DB2 family. The core capabilities of a distributed data server with respect to replication are shown in the top-left and top-right portions of Figure 1-12.

When you buy WebSphere Replication Server, you also get Q-based replication (Figure 1-13). In Q-based replication, each message represents a transaction that is placed on a WebSphere MQ Series message queue. This replication is known for its highly parallel apply process and rich conflict detection and resolution services.

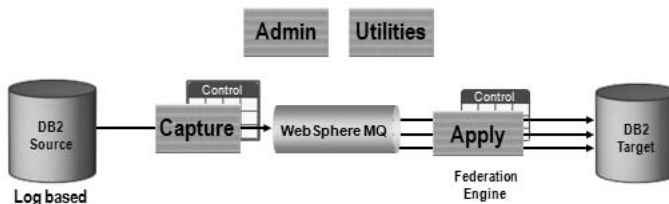


Figure 1-13 *The power of Q-based replication*

WebSphere Federation Server and WebSphere Information Server

WebSphere Information Server provides features and functions to place, publish, integrate, find, cleanse, and access data. The *federation* capabilities are worth mentioning as they relate to data access. WebSphere Federation Server contains a subset of the functionality found in WebSphere Information Server, namely the federation capabilities (and hence will be the focus of this section). You can learn more about WebSphere Federation Server and WebSphere Information Server at www.ibm.com/software/data/integration/.

A federated system is a distributed database environment that appears as a single virtual database to end users and applications. The WebSphere Federation Server technology (some of which is built into DB2) allows SQL statements to transparently access, join, or update data located across heterogeneous data sources. Federated support enables the DB2 data server to evaluate global statistics on all data sources in order to optimize requests for the remote data. For example, the global optimizer can use knowledge of source statistics, source indexes, source functions, server and network speeds, and so on.

The SQL support in a DB2 federated environment supports the same SQL as a non-federated environment, including advanced object-relational SQL operations across the federated data sources, such as recursive SQL, common table expressions (CTEs), and more. If the remote data source doesn't have this capability, DB2 will compensate for the missing functions.

Perhaps the greatest benefit to a federated system is that developers only have to learn a single dialect of SQL; namely, the DB2 ANSI-compliant SQL API. If an external data source has different data types or non-standard SQL, DB2 will transparently compensate for missing functions and convert data types. For example, SQL Server 2005 has a `MONEY` data type whereas DB2 uses a `DECIMAL` data type to support currency. Without the WebSphere Federation Server technology, a developer would have to know the differences between these data servers, or cast them to like data types using OLE DB. With WebSphere Federation Server, developers just code as if they were always accessing DB2 data, and the rest is taken care of.

An example of federated access is shown in Figure 1–14.

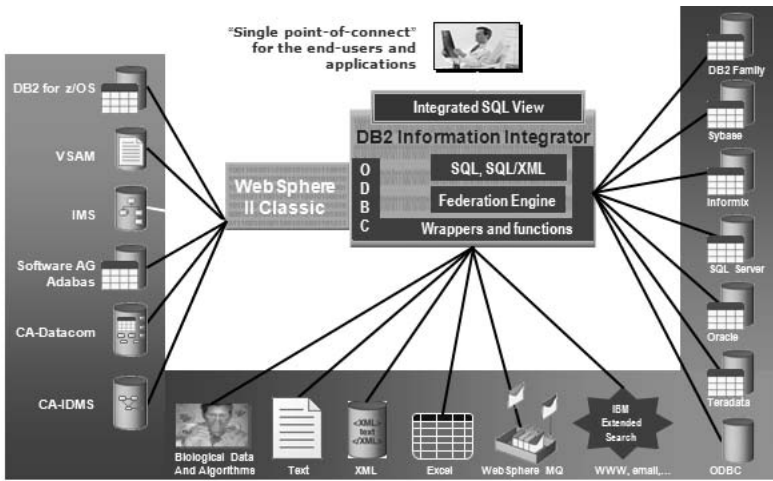


Figure 1-14 *The power of federation — any data, one API*

In Figure 1-15 you can see the technology components that make up a federated database. *Nicknames* are used to implement a low form of granularity with respect to the data sources you want to make available. Nicknames essentially are local aliases on remote tables, although they can be mapped to specific rows and columns. A collection of nicknames is located within a *server* that represents that actual data source. For example, in Figure 1-15 Oracle is one such data source, so the server component would represent a specific Oracle database. All the SQL eventually flows through a *wrapper*, which is a library of code that allows access to a particular class of data servers. Connections made to these servers use their native protocol; for example, the Oracle connection will use the Net8 client.

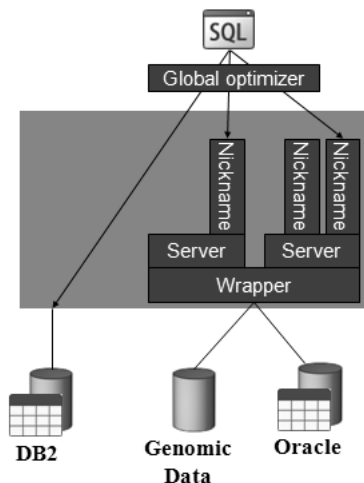


Figure 1-15 *How federation works*

You can see in Figure 1–14 that with WebSphere Federation Server you can pretty much extend the reach of DB2 to *any* data source in your environment. For example, in the life sciences industry, scientists need access to specialized data to support their research related to drug discovery. IBM offers a set of Life Sciences wrappers that can be used to build a federated system that can access data useful to scientists, including Excel spreadsheets, image data, flat files, and BLAST (Basic Local Alignment Search Tool).

If a wrapper is not available for purchase, there is a software developer's kit (SDK) that you can use to build your own.

Database Enterprise Developer's Edition

IBM offers a special discounted suite of Information Management products that you can purchase at a discount for development, evaluation, demonstration, and testing of your application programs; this edition is called Data Enterprise Developer Edition (DEDE).

Some of the products included in this comprehensive developer offering are:

- DB2 Express 9
- DB2 Workgroup 9
- DB2 Enterprise 9
- DB2 Runtime Client (including .msi merge modules on Windows)
- DB2 Client
- IBM Driver for ODBC and CLI
- IBM Driver for JDBC and SQLJ
- DB2 Developer Workbench
- DB2 Embedded Application Server (components of WebSphere Application Server)
- DB2 Information Center
- DB2 Documentation CD
- DB2e Enterprise Edition
- WebSphere MQ (this is a restricted copy that can only be used in conjunction with the DB2 software)
- Rational Web Developer (this is a restricted copy that can only be used in conjunction with the DB2 software)
- Tivoli System Automation (TSA) for Linux and AIX
- Informix IDS Enterprise Edition
- IBM Cloudscape/Apache Derby
- DB2 Connect Unlimited Edition for i5/OS and z/OS

- All the DB2 Extenders™, namely DB2 Spatial Extender, DB2 Geodetic Extender, DB2 Net Search Extender, and the DB2 XML Extender
- All of the DB2 add-on feature packs outlined earlier in this chapter

Most of the products within DEDE are available for all the platforms that DB2 Enterprise supports (unless of course a product doesn't exist on a specific platform). The data server where you install DEDE can be on a platform that is different from the one on which the application will eventually be deployed or tested because of the common code base used in the distributed DB2 environment. In fact, almost any applications developed using DEDE can be executed on any system that has DB2 client software (or specific drivers for which the application is written) installed.

The application development environment provided with DEDE allows application developers to write programs using today's most popular methodologies, including:

- Embedded SQL
- Call Level Interface (CLI)/Open Database Connectivity (ODBC)
- .NET Framework 2.0
- DB2 Application Programming Interfaces (APIs)
- Web Services
- Java Database Connectivity (JDBC) and SQLJ
- Python™
- PHP
- Perl™
- Ruby on Rails™

DEDE also includes the necessary programming libraries, header files, code samples, and pre-compilers for all of the supported programming languages.

DB2 Developer Workbench

The DB2 9 Developer Workbench (DB2 DWB) replaces the DB2 8 Development Center. The DB2 DWB is an Eclipse-based graphical environment that supports the rapid development of DB2 SQL and Java stored procedures, SQL scalar and table user defined functions (UDFs), SQL statements, XQuery statements, and Web Services. However, there's so much more to this list. For example, the DB2 DWB includes an SQL editor that's enriched with syntax colorization and code assistants, as well as teaming support, compare utilities, and more.

The DB2 Extenders encapsulate the attributes, structure, and behavior of these unstructured data types and stores this information in DB2. From the developer's perspective, the DB2 Extenders appear as seamless extensions to the database and enable the development of multimedia-based applications. In other words, a spatial data type is no different than a built-in data type that they may be accustomed to. This section briefly details the DB2 Extenders that are provided by IBM.

DB2 Spatial Extender

The DB2 Spatial Extender (DB2 SE) provides the ability to create spatially aware data objects and store them within your DB2 database, along with other spatially related objects like (LAT/LONG) coordinates and more. Almost all industries could benefit from this free technology in DB2. For example, the banking and finance industry could visually envelope customer segments for brand location identification. Municipal governments could use this technology for flood plain identification, the retail industry for billboard locations, and more. This seems apparent when you consider that almost all data has some sort of spatial component to it: we all have an address, merchandise in a warehouse has a stock location, and so on.

The business benefit of the DB2 SE lies in the notion that it's a lot easier to spot visually represented information than data reported in rows and columns.

When you enable your DB2 database for the DB2 SE, you can interact with your data using SQL or specialized spatial tools from other vendors. The point is that with the DB2 SE, DB2 understands the spatial "dialect" and the operations that you want to perform with it.

For example, a telematics application on a PDA may provide its users with a list of nearby Chinese restaurants that serve Peking Duck based on the dynamic request of this user. In this case, after the client's PDA creates a location box using Global Positioning System (GPS) coordinates, it could generate SQL statements similar to the following:

```
SELECT NAME, DESCRIPTION, ADDRESS FROM RESTAURANTS
WHERE
  OVERLAPS (LOCATION, BOX(GETGPS(), 2000, 2000))
  AND CATEGORY = 'CHINESE'
  AND DOC CONTAINS(MENU, 'PEKING DUCK');
```

OVERLAPS is a spatial function that shows interested data in a bounding box defined by the OVERLAPS boundary specification; there are many other spatial functions, including INTERSECTS, WITHIN, BUFFERS, and so on.

DB2 Geodetic Extender

The DB2 Geodetic Extender builds upon capabilities available in the DB2 Spatial Extender and adds compensation for real-world objects like the curvature of the earth's surface. The algorithms in this extender seek to remove the inaccuracies introduced by projections and so on. This extender is available only for DB2 Enterprise as part of the Data Geodetic Management feature.

DB2 Net Search Extender

The DB2 Net Search Extender (DB2 NSE) combines in-memory database technology with text search semantics for high-speed text search in DB2 databases. Searching with it can be particularly advantageous in Internet applications where performance is an important factor. The DB2 NSE can add the power of fast full-text retrieval to your DB2 applications. Its features let you store unstructured text documents of up to 2 GB in databases. It offers application developers a fast, versatile, and intelligent method of searching through such documents.

Additionally, the DB2 NSE provides a rich set of XML searching capabilities with advanced search features like sounds-like, stemming, and so on. It is shipped free in DB2 9 (it was a chargeable extender in DB2 8) to facilitate non-XML index searching of XML data stored in pureXML columns.

DB2 XML Extender

The DB2 XML Extender is provided with DB2 and allows you to store XML documents in DB2; it also gives you the ability to shred and store XML in its component parts as columns in multiple tables. In either case, indexes can be defined over the elements or attributes of an XML document for fast retrieval. Furthermore, text and fragment search can be enabled on the XML column or its decomposed parts via the DB2 Net Search Extender. The DB2 XML Extender can also help you formulate an XML document from existing DB2 tables for data interchange in business-to-business environments.

You may recall that the pureXML add-on feature pack is available for all DB2 9 data servers. Indeed, this can cause confusion since the DB2 XML Extender is shipped for free in DB2 9. You should consider the DB2 XML Extender as stabilized technology. In other words, it is no longer being enhanced and shouldn't be considered for most XML applications. The DB2 XML Extender's approach to storing XML is to shred the XML to relational tables or stuff it into a large object. When you use this technology to persist XML data, you have to make serious trade-offs with respect to performance, flexibility, and so on. In addition, you have to use specialized functions to implement Spathe searches, and data types are abstracted from base DB2 data types. Quite simply, the way you interact with the DB2 XML Extender isn't natural for XML programmers and DBAs alike.

In contrast, the pureXML feature in DB2 9 provides services such that no compromises between flexibility (what XML was designed for) and performance (one of the reasons why you want the data server to store your XML) need to be made when storing your XML data. For example, to generate XML documents from relational tables, you simply use the SQL/XML API instead of the cumbersome DB2 XML Extender functions. You can validate XML documents against Sods instead of only document type definitions (Ds) as is the case with the DB2 XML Extender, and more. We strongly recommend this feature for most of your XML-based applications.

DB2 Administration

DB2 DBAs have a number of graphical-based tools they can use to manage and administer DB2 data servers. Alternatively, a DBA can also use a script-based approach to administer the data environment using the DB2 tools to create and schedule the scripts. This section briefly details the main graphical tools available with DB2.

Control Center

The Control Center is the central point of administration for DB2. The Control Center provides DBAs with the tools necessary to perform typical database administration tasks. It allows easy access to other server administration tools, gives a clear overview of the entire system, enables remote database management, and provides step-by-step assistance for complex tasks.

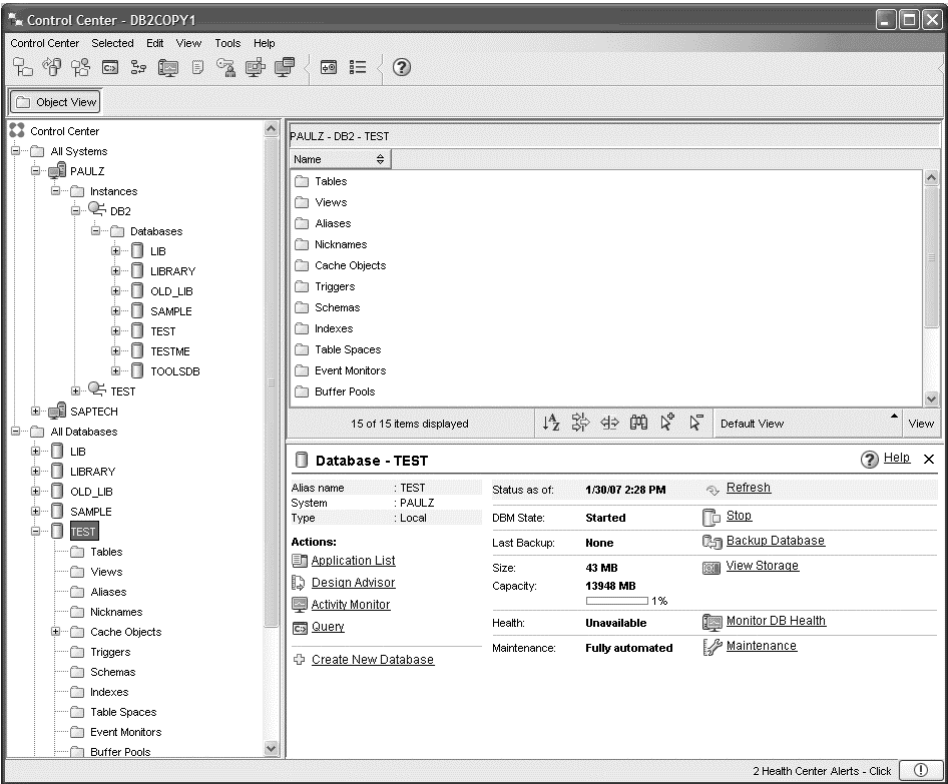


Figure 1–17 The DB2 Control Center

The **All Systems** object represents both local and remote data servers. To display all the DB2 systems that your system knows about, expand the object tree by clicking on the plus sign (+) next to **All Systems**. In Figure 1–17, you can see a DB2 data server called PAULZ contains a DB2 instance called DB2, in which the database TEST is located.

When you highlight an object, details about that object are shown in the Contents Pane.

The main components of the Control Center are:

- **Menu Bar** — Used to access Control Center functions and online help.
- **Tool Bar** — Used to access other DB2 administration tools, such as the Command Editor, Task Center, and more.
- **Objects Pane** — This is shown on the left side of the Control Center window. It contains all the objects that can be managed from the Control Center as well as their relationship to each other.
- **Contents Pane** — This is found on the right side of the Control Center window and contains the objects that belong or correspond to the object selected in the Objects Pane.
- **Contents Pane Toolbar** — These icons are used to tailor the view of the objects and information in the Contents pane. These functions can also be selected in the View menu.
- **Task Window** — Lists the most common tasks associated with the selected object in the Object Pane. In Figure 1–17 you can see that since a database is highlighted, common tasks and administrative functions related to it are in this window.
- **Hover Help** — Provides a short description for each icon on the toolbar as you move the mouse pointer over the icon.

The Control Center also comes with *personality control* that you can use to adjust the view and functions available from the Control Center's tree view of your data server. For example, you can limit the Object Pain view to show just Tables or Views, as well as limit the actions you can perform from the context-sensitive right-click menu options. You can customize your Control Center personalities using **Tools → Tools Settings → Customize Control Center**.

DB2 Note

The facility to define a Control Center personality by defaults pops up each and every time you start the Control Center. You can turn off this option by deselecting the **Show this window at startup time** checkbox.

DB2 Replication Center

The DB2 Replication Center is a graphical tool that allows DBAs to quickly set up and administer all forms of data replication, including the options offered by WebSphere Replication Server. The main functions in setting up a replication environment can be performed with this tool, including:

- Registering replication sources
- Monitoring the replication process
- Operating the CAPTURE and APPLY programs
- Defining alerts

You can use the Replication Center to set up all kinds of DB2 replications, as shown in Figure 1–18.

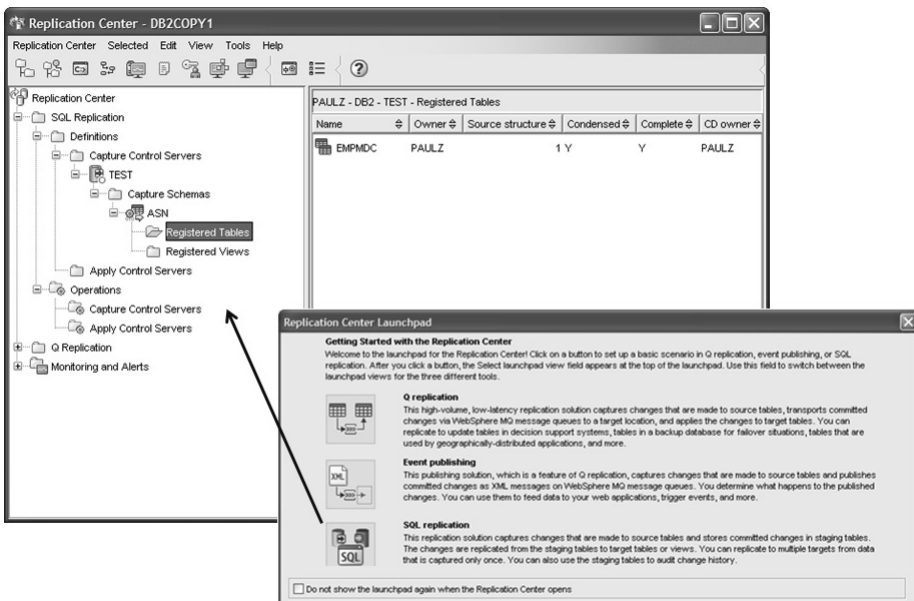


Figure 1–18 *The DB2 Replication Center*

Other Tools Available from the Control Center

By using the Control Center tool bar, you can access a host of other graphical administration tools to help you manage and administer databases in your environment:

- **Satellite Administration Center** — Used to manage groups of DB2 data servers through push/pull management scripts and more.
- **Command Editor** — Provides an interactive window that facilitates the building of SQL statements or DB2 commands, the viewing of execution results, and explain information. This graphical command utility is often the preferred method for text commands as it provides enormous flexibility and functionality.
- **Task Center** — Used to create, schedule, and manage scripts that can contain SQL statements, DB2 commands, or operating systems commands.
- **Journal** — Keeps a record of all script invocations, all DB2 messages, and the DB2 recovery history file for a database. It is used to show the results of a job, to display the contents of a script, and also to enable or disable scheduled jobs.
- **License Center** — Used to manage licenses and check how many connections are used.
- **DB2 Web Tools** — Allows the DBAs to use an HTTP client to remotely execute SQL statements, DB2 commands, or operating system commands against a DB2 server. Essentially, the Health Center and the Command Editor are exposed in this tool set.

DB2 Health Center

The DB2 Health Center (Figure 1–19) is the central point of information with respect to managing the health of your DB2 system. When you install a DB2 9 data server, out of the box it automatically monitors 27 (and counting) health indicators that proactively monitor the health of your data server. The DB2 Health Center implements a management-by-exception model whereby alerts are surfaced when Warning or Alarm thresholds are breached. Although DB2 configures these thresholds for you out of the box, you can configure them yourself as well as specify scripted actions to occur in the event of an alert.

You don't have to use the DB2 Health Center to work with DB2's health information or set triggered actions to occur on threshold breaches. You can use a number of SQL-based user defined functions to work with the DB2 health facilities from the command line.

The DB2 Health Center can monitor indicators across the following health categories:

- Application concurrency (e.g., Deadlock rate)

- DBMS (e.g., Instance operational state)
- Database (e.g., Database operational state)
- Logging (e.g., Log utilization)
- Memory (e.g., Monitor heap utilization)
- Sorting (e.g., Percentage of sorts that overflowed)
- Table space storage (e.g., Table space utilization)
- Database maintenance (e.g., Reorganization required)
- High-Availability Disaster Recovery (e.g., HADR log delay)
- Federated (e.g., Nickname status)
- Package and catalog caches (e.g., Package cache hit ratio)

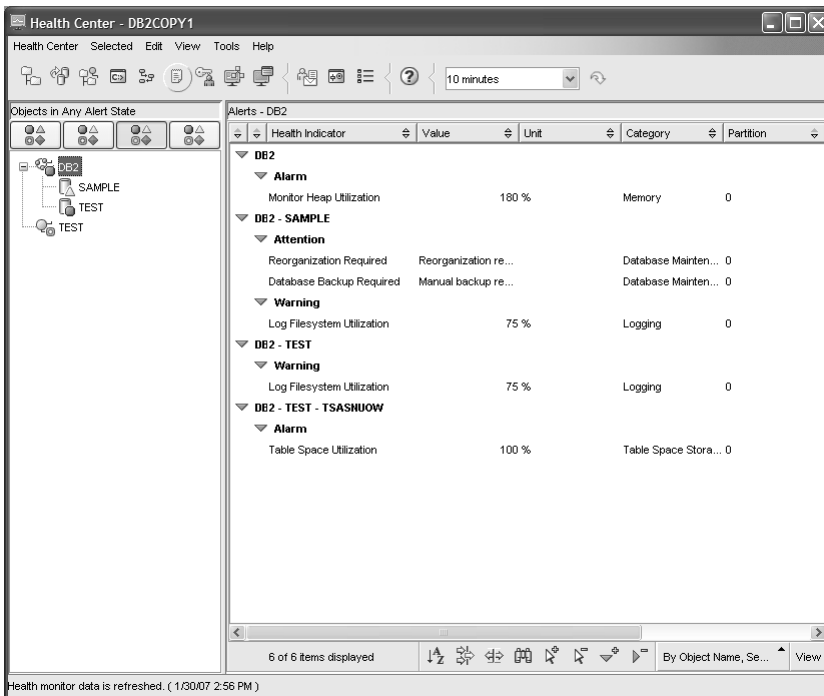


Figure 1–19 *The DB2 Health Center*

The DB2 Health Center's graphical user interface allows DBAs to select database objects and drill down on its details, current alerts, and the recommended actions. The DB2 Health Center also includes the DB2 Recommendation Advisor that can be used to walk through potential fixes to alerts and alarms raised in this facility.

DB2 Configuration Assistant

The DB2 Configuration Assistant (DB2 CA) lets you maintain a list of databases to which your applications can connect, as well as manage and administer those connections. It is mostly used for client connectivity configuration. You can start the DB2 CA by entering the `db2ca` command from your operating system's command-line shell or from the DB2 folder in the **Start** menu. Some of the functions available in the DB2 CA are shown in Figure 1–20.

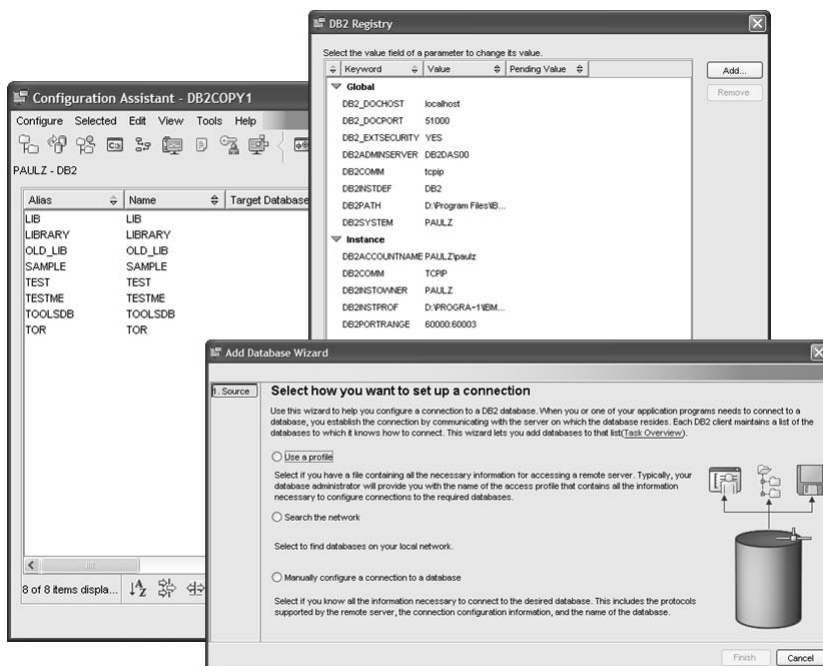


Figure 1–20 *The DB2 Configuration Assistant*

Using the DB2 CA, you can work with existing databases, add new ones, bind applications, set client configuration and registry parameters (also shown in Figure 1–20), test connectivity, and import and export configuration profiles.

The DB2 CA's graphical interface makes these complex tasks easier by means of the following:

- Wizards that help you perform certain tasks
- Dynamic fields that are activated based on your input choices
- Hints that help you make configuration decisions
- The Discovery feature, which can retrieve information that is known about databases that reside on your network

The DB2 CA's Discovery feature is very useful because it allows you to add a database connection without having to know the syntax of DB2 CATALOG NODE and DB2 CATALOG DATABASE commands, or even the location information of the remote data server.

As you can see in Figure 1–20, the DB2 CA displays a list of the databases to which your applications can connect from the workstation where it was started. Each database is identified first by its database alias, then by its name. You can use the Change Database Wizard to alter the information associated with databases in this list. The CA also has an Advanced view, which uses a notebook to organize connection information by the following objects:

- Systems
- Instance nodes
- Databases
- Database Connection Services (DCS) for System i and System z databases
- Data sources

Advisors and Wizards

DB2 comes with a set of Wizards and Advisors to help you with day-to-day tasks. Wizards can be very useful to both novice and expert DB2 users. Wizards help you complete specific tasks by taking you through each task one step at a time and recommending settings where applicable. Wizards are available through both the Control Center and the Configuration Assistant.

There are wizards for adding a database to your system, creating a database, backing up and restoring a database, creating tables, creating table spaces, configuring two-phase commit environments, configuring database logging, updating your documentation, setting up a High-Availability Disaster Recovery (HADR) pair, tuning your performance, and more.

Figure 1–21 shows a portion of the Create Database wizard in DB2 9.

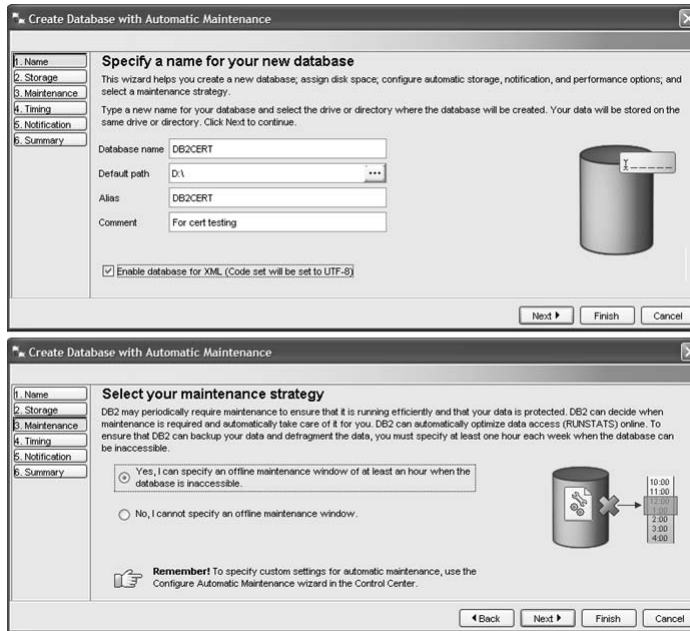


Figure 1–21 *The Create Database wizard*

Advisors are special types of wizards that do more than provide assistance in completing a task. Traditional wizards take you step-by-step through a task, simplifying the experience by asking important questions or generating the complex command syntax for the action you want to perform. When a wizard has more intelligence than just task completion and can offer advisory-type functions, DB2 calls them *advisors*. They operate just like wizards but are intelligent enough (having some pretty complex algorithms) to generate advice based on input factors such as workload or statistics. Advisors help you with more complex activities, such as tuning tasks, by gathering information and recommending options that you may not have considered. You can then accept or reject the advisor's advice. You can call advisors from context menus in the DB2 administration tools, from APIs, and the command-line interface.

Advisors are part of the IBM autonomic computing effort, which aims to make software and hardware more SMART (self-managing and resource tuning). There are three main advisors in DB2 9: the DB2 Configuration Advisor, the DB2 Recommendation Advisor, and the DB2 Design Advisor.

The DB2 Configuration Advisor is automatically run for you whenever you create a new database in DB2 9. It can configure up to 35 instance-level and database-level parameters for you based on responses to high-level questions that describe the data server environment and type of application you plan to support.

The DB2 Recommendation Advisor, as previously mentioned, is closely associated with the DB2 Health Center and is used to offer solutions to raised alerts and alarm breeches in this facility.

The DB2 Design Advisor is used to identify objects such as materialized query tables (MQTs), multidimensional clustering tables (MDCs), indexes, and partitioning keys that could optimize a given SQL workload. The DB2 Design Advisor can also identify indexes that aren't needed as well (shown in Figure 1–22).

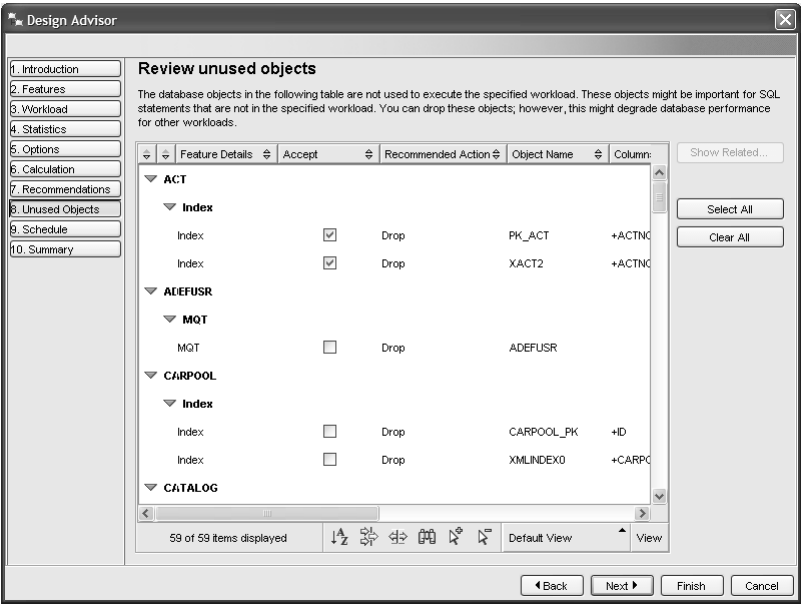


Figure 1–22 The DB2 Design Advisor

When using this advisor it's important to note that the suggestions it provides are based on a submitted workload. If you've left out significant portions of a workload, the answer will not reflect the impact of the missing workload. In addition, the DB2 Design Advisor gives you the ability to heavily weight SQL statements in a submitted workload over others, giving you more control with respect to how the DB2 Design Advisor will recommend the creation of performance objects with respect to your real workload characteristics.

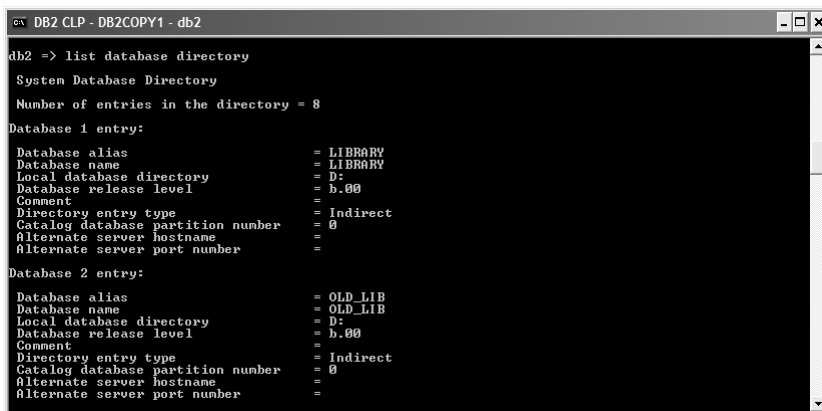
The DB2 Command Line Processor

The DB2 Command Line Processor (DB2 CLP) is a component common to all DB2 products. It is a text-based application that can be used to issue interactive SQL statements or DB2 commands. For example, you can create a database, catalog a database, and issue dynamic SQL statements all from the DB2 CLP. Your DB2 statements and commands can also be placed in a file and executed in a batch environment, or they can be entered in interactive mode.

DB2 Note

Commands issued from the DB2 CLP can also be issued from the DB2 Command Editor or from most operating system's command line processor (CLP) using the `db2` prefix.

Figure 1–23 shows an example of using the DB2 CLP to enter DB2 commands. The DB2 CLP operates in an “interactive” mode and therefore does not require the `db2` prefix associated with entering DB2 commands.



```
DB2 CLP - DB2COPY1 - db2
db2 => list database directory

System Database Directory
Number of entries in the directory = 8

Database 1 entry:
Database alias           = LIBRARY
Database name            = LIBRARY
Local database directory = D:
Database release level   = b.00
Comment                  =
Directory entry type     = Indirect
Catalog database partition number = 0
Alternate server hostname =
Alternate server port number =

Database 2 entry:
Database alias           = OLD_LIB
Database name            = OLD_LIB
Local database directory = D:
Database release level   = b.00
Comment                  =
Directory entry type     = Indirect
Catalog database partition number = 0
Alternate server hostname =
Alternate server port number =
```

Figure 1–23 *Using the DB2 CLP to enter commands*

The DB2 CLP is provided with all DB2 and DB2 Connect products. All SQL statements issued from the DB2 CLP are dynamically prepared and executed on the data server. The output, or result, of the SQL query is displayed on the screen by default. All of the DB2 commands that you can enter in the DB2 CLP are documented in the *DB2 Command Reference*. You learn more about the DB2 CLP and how to enter DB2 commands from an operating system's native CLP in Chapter 2.

Visual Explain

Other graphical tools can be used for tuning or monitoring performance. Visual Explain is a graphical utility that provides a visual representation of the access plan that DB2 uses to execute an SQL statement.

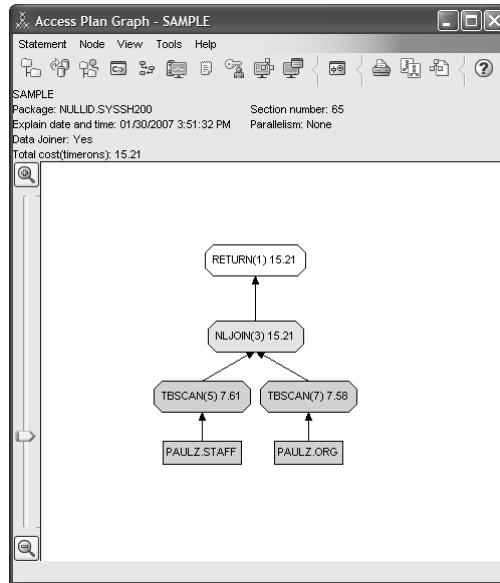


Figure 1–24 Using Visual Explain to look at how your query is run by DB2

Visual Explain can be invoked from the Control Center, the DB2 CLP (though the output in textual and not graphical), the DB2 DWB, or from the Command Editor.

Figure 1–24 shows the type of information that is displayed. You can see that the query is accessing two tables and an approximation of the cost of each step of this query is also provided in the Visual Explain output. The estimated query costs represent the complexity and resource usage expected for a given SQL query. There are many more details and features provided in Visual Explain under the **Statement** menu option.

DB2 Query Patroller

DB2 Query Patroller (DB2 QP) is an add-on product that can be used to control and monitor query execution, as well as work with queries to prioritize and schedule user queries based on user profiles and cost analysis performed on each query. Large queries can be put on hold and scheduled for a later time during off-peak

hours. Queries with high priority (based on user profiles) are promoted to the top of the schedule.

In addition, DB2 QP monitors resource utilization statistics. DB2 QP can use this information to determine the load distribution of the system, which can allow it to balance the number of users allowed to submit queries at any given time.

DB2 QP greatly improves the scalability of a data warehouse by allowing hundreds of users to safely submit queries on multi-terabyte class data servers. Its components span the distributed environment to better manage and control all aspects of query submission. The services provided by this product act as an agent on behalf of the end user. It prioritizes and schedules queries so that query completion is more predictable and system resources are more efficiently utilized. DB2 QP obtains query costs from the DB2 Optimizer and then schedules them for execution — this means that DB2 QP is tightly integrated with DB2 engine.

DB2 QP can also be used to set individual user and user class priorities as well as user query limits. This enables the data warehouse to deliver the needed results to its most important users as quickly as possible. If desired, an end user can choose to receive notice of scheduled query completion through e-mail.

Finally, as discussed earlier in this chapter, DB2 QP offers the ability to perform charge back for data server usage to specific departments identified by accounting strings on the connection context. For example, if marketing is using the data warehouse three times more than accounting, they should pay three times the charge back. In Figure 1–25, you can see one of many canned reports that come with DB2 QP. This one shows the number of statements run by month. You can drill down into this view to the minute and second interval, as well as access a host of other reports such as average execution time, average wait time, average queue time, and more.

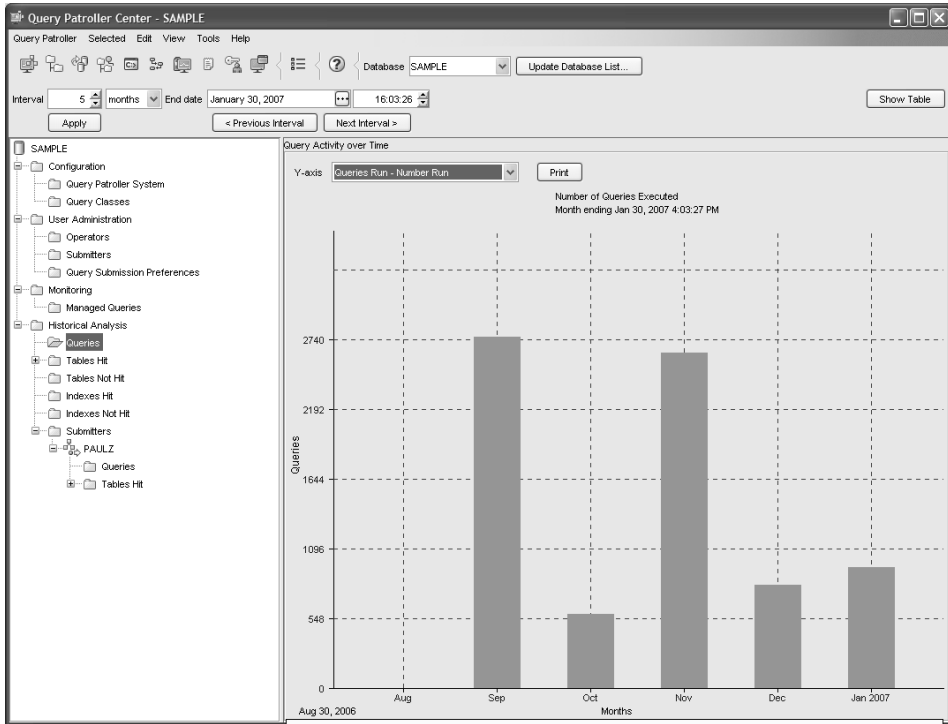


Figure 1–25 Performing charge back with DB2 Query Patroller

Database Monitoring Tools

The *Snapshot Monitor* captures database information at specific intervals. The interval time and data represented in the performance graph can be configured. The Snapshot Monitor can help analyze performance problems, tune SQL statements, and identify exception conditions based on limits or thresholds.

The *Event Monitor* captures database activity events as defined by the event monitor definition. Event Monitor records are usually stored on disk and then analyzed after the data has been captured. The *Event Analyzer* graphical tool provided with DB2 can be used to analyze the captured data.

The *Activity Monitor* help you improve the efficiency of database performance monitoring, problem determination, and resolution. By tracking a set of predefined monitor data, the Activity Monitor allows you to quickly locate the cause of a problem. You can then take direct action to resolve the problem or invoke another tool for further investigation. The Activity Monitor can help you monitor application performance, application concurrency, resource consumption, and SQL statement

usage. It can also assist you in diagnosing performance problems such as lock waiting situations (as shown in Figure 1–26), and in tuning queries for optimal utilization of the data server’s resources.

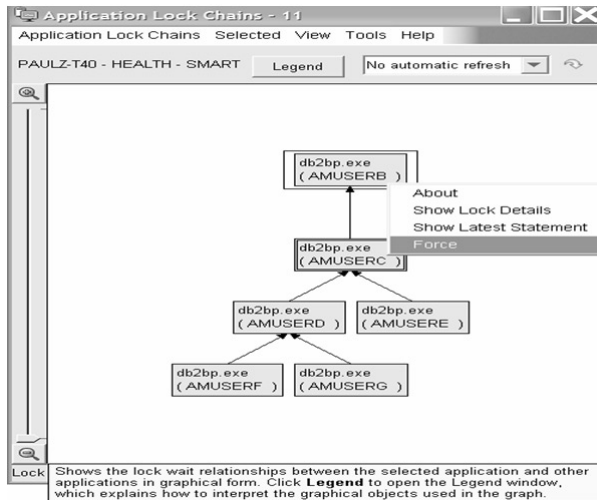


Figure 1–26 *Using the Activity Monitor to diagnose a lock waiting problem*

The DB2 Storage Management Tool

The DB2 Storage Management Tool can be used to monitor the storage state of a database. You can use this facility to take storage snapshots for a database or a table space. When a database or snapshot is taken, statistical information is collected for all the table spaces defined in the given database (you can also snapshot at the table space level).

The Storage Management Tool enables you to set thresholds for data skew, space usage, and index cluster ratio. If a target object exceeds a specified threshold, the icons beside the object and its parent object in the Storage Management view are marked with a warning flag or an alarm flag — similar to the DB2 Health Center.

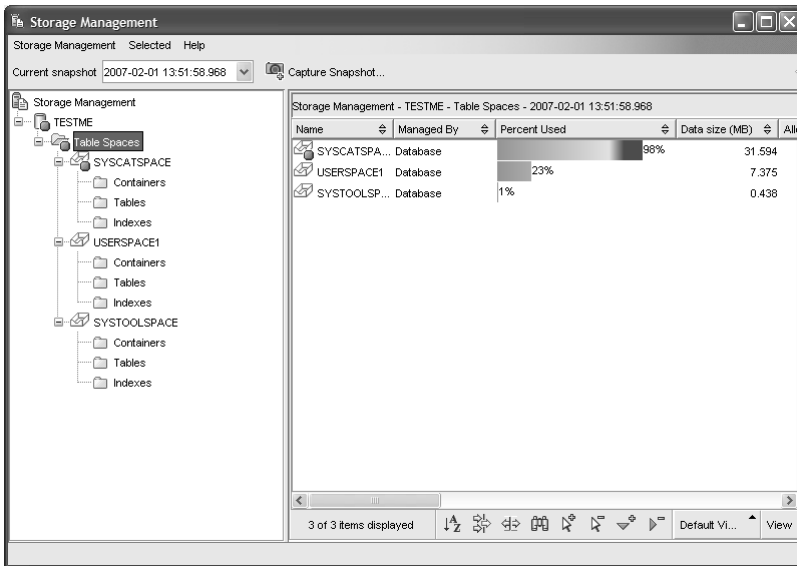


Figure 1–27 Using the DB2 tools to manage storage activity

You can see in Figure 1–27 that the SYSCATSPACE table space is running out of space as it's 98 percent used.

The DB2 Aftermarket Tools

There are two kinds of tools for DB2: those that are free and those that are add-ons that can be purchased separately. The free tools come as part of a DB2 installation and can be launched from the Control Center, the Configuration Assistant, or on their own. A separate set of purchasable tools are available to help ease a DBA's task of managing and recovering data, tuning performance, and more. The DB2 suite of these tools includes (www-306.ibm.com/software/data/tools/mptools.html):

- **DB2 Change Management Expert** — Improves DBA productivity and reduces human error by automating and managing complex DB2 structural changes.
- **Data Archive Expert** — Responds to legislative requirements like Sarbanes-Oxley by helping DBAs move seldom-used data to a less costly storage medium without additional programming.
- **DB2 High Performance Unload** — Maximizes DBA productivity by reducing maintenance windows for data unloading and repartitioning.
- **DB2 Performance Expert** — Makes DBAs more proactive in performance management to maximize database performance. (This tool was discussed in the

“Add-on Feature Packs for DB2 Enterprise Edition” section earlier in this chapter).

- **DB2 Recovery Expert** — Protects your data by providing quick and precise recovery capabilities, including operations only provided in this tool like SQL statement undo generation, object recovery, and more.
- **DB2 Table Editor** — Keeps business data current by letting end users easily and securely create, read, update, and delete (CRUD) data.
- **DB2 Test Database Generator** — Quickly creates test data and helps avoid liabilities associated with data privacy laws by protecting sensitive production data used in test.
- **DB2 Web Query Tool** — Broadens end user access to DB2 data using the Web and handheld devices.

Summary

This chapter discussed the DB2 products for LUW. There are a number of offerings available, including:

- DB2 Enterprise Edition
- DB2 Workgroup Edition
- DB2 Express Edition
- DB2 Express-C
- DB2 Personal Edition
- DB2 Everyplace Edition
- Various DB2 Connect Editions
- DB2 Query Patroller

These products provide the flexibility to execute database applications running on pervasive devices up to multi-node clusters. DB2 provides support for the commonly used communication protocols.

Each of the DB2 data server editions comes with a set of purchasable add-on feature packs that you can use to extend the capabilities of the core data server. There are also number of add-on tools and products that you can also buy for DB2, including DB2 Query Patroller and a myriad of DBA-focused tools such as High Performance Unload, the DB2 Recovery Expert, and more.

SQL-based replication is integrated into all DB2 for Linux, UNIX, and Windows servers. Replication can be used to move data between members of the DB2 family, or from and to members of the DB2 family and non-DB2 data servers (like Oracle, SQL Server, and more) if you’re using WebSphere Replication Edition. DB2 Connect is used to access DB2 data that resides on DB2 for i5/OS, VM/VSE, and

z/OS operating systems, and it comes with federated capabilities that can also be added to a DB2 data server installation.

DB2 9 includes a number of application development enhancements, including the DB2 Developer Workbench and integration into the world's most popular IDEs such as Rational Application Developer, Zend Core, and Microsoft Visual Studio 2005.

Significant new features have been added to make DB2 easier to manage. SMART technology (Self-managing and Resource Tuning) has been integrated into a number of DB2 components, including installation, configuration, utilities, problem determination and resolution, and availability. This is part of IBM's autonomic computing initiative and new features will continue to be added to make DBAs more productive.

This chapter also introduced some of the graphical and command line tools available in DB2. The Command Line Processor (CLP) is a text-based application that allows you to enter DB2 commands and SQL statements and is found in all DB2 products. From the desktop, an administrator can configure remote and local systems, administer instances and databases, and create database objects graphically using the Control Center or the Configuration Assistant. Tools like the DB2 Health Center and DB2 Activity Monitor also help manage DB2 environments while export support is provided in the form of Wizards and Advisors. In the remaining chapters, additional DB2 functions and tools will be examined for how they assist the end user, application developer, and administrator.