



Attunity Connect
software



a white paper
from hp

middleware for universal data access

The media often talk about today's tidal waves of business information—massive and, by implication, overwhelming. What this compelling image doesn't convey is that "massive" is not necessarily the problem. The diversity of today's data—pouring in from the Internet, created by different software tools, hosted on numerous internal and external platforms, residing in a multitude of legacy databases—is what's truly overwhelming.

This diversity in the structure and complexity of data is driving the need for Universal Data Access middleware capable of providing businesses with a standard, unified view of data, whether it originates within or outside an organization. Attunity Connect software is a new-generation, standards-based middleware solution from Attunity, Inc. With this multi-tiered data access and integration solution, HP NonStop™ system users can access and manipulate data from diverse data sources, residing on both local and remote platforms and across the Internet and intranets. In providing this capability, Attunity Connect software opens the door to standard desktop applications and tools, server-based applications, and new classes of Web applications that can take advantage of the massive amounts of information now theoretically available but too often locked up in incompatible systems.

universal data access

With the increased capacity of PC software and the extended availability of external and remote data, businesses are discovering the need for solutions that span desktop, midrange, legacy mainframe, and Internet technologies. Today, as little as 20 percent of a typical organization's data resides internally. The remainder usually comes from outside sources, including the Internet. In addition, much of the data in use today is legacy data, stored on different systems and in divergent formats.

Data access middleware enables an application to access data from a remote source, regardless of location and format. It does this by buffering the application from the native interface of the data source, the location of the data source, and the data model. Middleware for Universal Data Access takes this a step further by presenting a single, consistent interface and data model (along with multiple data drivers) for many different types of data sources on many different platforms.

industry standards

The Open Database Connectivity (ODBC) standard enables universal access to relational data in a point-to-point format. The introduction of the OLE DB and Microsoft® ActiveX Data Objects (ADO) models then made possible the construction of industry-standard Universal Data Access and join capabilities for all data—regardless of type, format, or location.

transparent access to data from multiple sources

Universal Data Access and its supporting technologies allow businesses to draw data from practically any source, inside or outside the organization. Data access middleware processes application requests for data from a variety of database types—first packaging application requests for data and then transporting them to a specific server system, which handles the request. After processing the requested data, the middleware returns the data to the application.

One example is an application for risk management. A risk manager could tap data on a regular basis from internal departmental databases as well as local spreadsheet programs, while also pulling in data from industry groups and governmental and public databases via the Internet or private networks. As another example, a customer care or call center application could access data from various legacy sources—including customer databases and order entry databases.

In these and similar cases, the format and location of the data is completely masked by the data access middleware and is thus transparent to the application. The user must be concerned only with the information needed, as well as when and where it should be delivered.

requirements for universal data access middleware

To deliver on the promise of Universal Data Access—information when and where you want it, not limited by location or divergent formats or platforms—data access middleware must meet several key requirements. These include the ability to

- Be distributed throughout an enterprise
- Reside on multiple platforms
- Address diverse data sources and formats
- Join different types of data from multiple sources
- Provide performance similar to direct access

multiple platforms and distributed architecture

Universal Data Access is an enterprise requirement. Data access middleware should support deployment on a single client or server, as well as distributed configurations for enterprisewide deployment. Just as data must be accessed from multiple platforms, the middleware must reside on the full range of enterprise platforms.

diverse data sources and formats

Data access middleware must present a consistent data model to standard application program interfaces (APIs), regardless of the multiplicity of back-end data sources. This allows applications to access all desired data regardless of storage format—from

simple files to relational database management systems (RDBMSs) to legacy databases. This capability should extend beyond SQL databases to encompass all types of data providers.

joining data from multiple sources

Often, data must be combined, or joined, from multiple local and remote sources. These sources are frequently in heterogeneous or nonrelational environments. The data access middleware must provide the functionality necessary for querying any native data source, even when the source itself does not have SQL querying capabilities.

high performance

All middleware places a tier of functionality between an application and source data or another application. Data access middleware must provide performance equal to that of the proprietary interface supplied by native data sources for direct access. In addition, queries must be optimized to minimize traffic on the network and reduce the use of processing resources on the servers.

true universal data access for hp customers

Attunity Connect software adds Universal Data Access capabilities to NonStop SQL software's advanced data mining, data warehousing, object handling, and query optimization features. Attunity Connect software also accesses the popular Enscribe file system from HP.

Fully compliant with the OLE DB specification, Attunity Connect software enables access to data sources such as hierarchical databases, indexed sequential files, simple files, personal databases, spreadsheets, mail folders, and other sources—all through a uniform interface. Nonrelational data can be accessed in the same manner as relational data. And data from different sources—relational or nonrelational, internal or external—can be integrated in a single query result.

HP customers now can create applications that are independent of database systems and transparent to the physical location of those systems. All of a company's databases can be turned into ADO/OLE DB, ODBC, or Java™ Database Connectivity (JDBC) resources that provide universal yet secure access for applications—including Internet and intranet applications—to a wealth of corporate information (see figure 1).

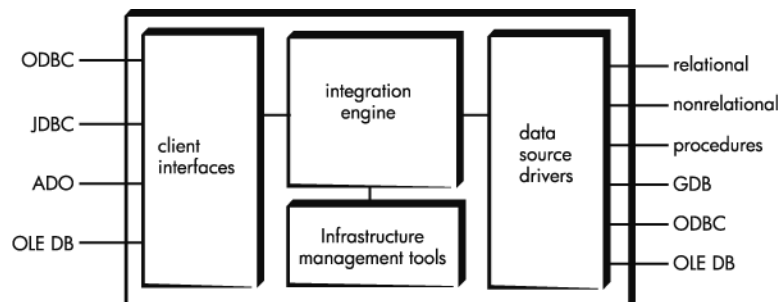


Figure 1. Attunity Connect software enables data from diverse sources to be accessed and combined for delivery to client applications.

enterprise deployment and universal access

Attunity Connect software runs identically on—and transparently accesses data from—multiple platforms. These include Microsoft Windows® 95 and 98, Windows NT® Server, and Windows 2000 operating systems running on ProLiant servers from HP; most popular versions of the UNIX® operating system, including HP Tru64 UNIX; HP OpenVMS operating system; IBM MVS; AS/400 (planned); and the NonStop platform. Applications written on any of these platforms can use Attunity Connect software to access data from any other supported platform. (See table.)

In addition, due to Attunity Connect software’s distributed object-oriented architecture, multi-tiered configurations combining various platforms can be deployed for enterprisewide data access.¹ When distributed over an enterprise network (using built-in TCP/IP links), the Attunity Connect components interact in a way that minimizes communications traffic.

Once deployed, Attunity Connect software enables data to be accessed and combined for delivery to client applications from relational databases such as NonStop SQL software, Oracle® databases, Microsoft SQL Server, DB2, Sybase, Informix, and relational OLE DB sources; from nonrelational databases including ADABAS and IMS DB (planned); and from file systems such as Enscribe, RMS, C-ISAM, D-ISAM, VSAM, text files, and flat files.

Attunity Connect software data access and platform support

<i>comprehensive data access</i>	<i>integration with business and development tools</i>	<i>full implementation of standards</i>	<i>extensive platform support</i>
NonStop SQL	Visual Basic	ADO	Windows 95 and 98, Windows NT Server, and Windows 2000 on Intel® processor-based systems, including ProLiant servers
Enscribe	C, C++, and Visual C++	OLE DB	
Oracle	Delphi	ODBC	
SQL Server	PowerBuilder	JDBC	
DB2	Business Objects	Oracle Heterogeneous Services Interface	
Sybase	Visual InterDev		
Rdb	Visual J		
RMS	SilverStream		
Oracle Codasyl DBMS	Internet tools such as		
VSAM	Internet Information Server (IIS), Active Server Pages, and RDS		
IDMS			OpenVMS operating system
C-ISAM, D-ISAM	Reporting tools like Crystal Reports and Impromptu		IBM MVS
Informix			NonStop systems
OpenIngres	Microsoft Query, Access, and Excel		
ADABAS			
Record-level OLE DB	Other ADO-, ODBC-, and JDBC-compliant tools		

¹ This software’s object orientation is based on Microsoft’s Component Object Model (COM). On platforms where COM and Distributed COM (DCOM) are not available, Attunity Connect software includes an object manager, which provides DCOM-like connectivity services to objects.

addressing diverse data sources

Attunity Connect software provides true and complete database transparency to its applications. This is due primarily to its OLE DB interface, which presents a consistent data model. OLE DB is an object-oriented architecture that extends the scope of data access beyond the relational model offered by ODBC, thus enabling access to any type of information regardless of format or storage method.

The OLE DB architecture includes components such as query engines, cursor engines, optimizers, business rules, and transaction managers. Because of this, Attunity Connect software allows data sources lacking one or more of these capabilities (such as file systems without query processing and cursor management) to benefit from this functionality. Similarly, Attunity Connect software provides the necessary SQL extensions to support hierarchical queries and queries over embedded rowsets. Its ODBC API, meanwhile, supports the hundreds of desktop applications relying on the ODBC standard.

Attunity Connect software also supports ADO, the object-based application-level interface to OLE DB. The successor to Remote Data Objects (RDO) and Data Access Objects (DAO), ADO provides a new common data access object for use across all applications and tools (including Visual Basic development system, C/C++, and Java tools). Microsoft products—such as Active Server Pages, which is used to develop scripts for Internet and intranet applications—use ADO for accessing data from OLE DB data providers. In conjunction with Active Server Pages and with native support for ADO, Attunity Connect software makes it easy to deploy intranet applications that make the full range of enterprise data accessible in an integral fashion via any Web browser.

Similarly, for those wishing to use Java tools and software, Attunity Connect software supports the JDBC interface. As a result, applications using JDBC requests benefit from all the Universal Data Access functionality of Attunity Connect software.

The software's ability to access any data source and its ability to join together data from multiple diverse sources enable the development of new types of applications. Users need to specify access details for each of their data sources only once. Different data sources can then be accessed via the same standard SQL query, simply by prefixing the desired data source tables or records with the data sources' respective names. It doesn't matter if some sources are relational and others are not.

powerful query processing

When joining data from diverse local and remote sources, Attunity Connect software employs a sophisticated distributed query processor and optimizer, designed to boost performance in heterogeneous and nonrelational environments. The query processor not only enables users to join data from multiple sources, but also compensates for functionality not available in the native data source, particularly through its support of SQL queries on data sources without query capabilities.² It also makes use of extended SQL so applications can access and manipulate hierarchical and other non-flat-file data while preserving the levels of detail needed for efficient drill-down operations.³

² Attunity Connect software includes drivers that allow access to nonrelational data (such as Oracle Codasyl DBMS, a Codasyl-type "network" system, or RMS and Enscribe indexed file systems). Each driver maps the specific functionality of its data source to OLE DB objects. The query processor accepts standard SQL syntax and restructures the SQL requests into optimized OLE DB objects.

³ Attunity Connect software provides the SQL extensions necessary to support access to hierarchically structured data (such as arrays). The hierarchical nature of the data is maintained in the output, allowing users either to arrange the output in a hierarchy or to flatten it via virtual tables.

Because of its distributed architecture, the Attunity Connect query processor and optimizer can perform joins and filtering as close as possible to the physical location of the data. This allows an application to use the processing power of the database servers to optimize application performance and also minimizes network traffic.

cost-based query optimizer

The cost-based query optimizer fulfills a critical performance role because many data sources do not have their own query processing or optimization capabilities. In addition, a heterogeneous query may access numerous different data sources and global optimization must be done outside the scope of any one of them.

Every query is parsed, analyzed, and optimized, after which an execution plan is formed. If a query can be sent to a relational database system (local or remote) for execution, this is always the selected approach: The optimizer assumes it cannot execute queries more efficiently than the native database system. If, however, tables from several RDBMSs or from nonrelational systems are involved, the query processor decides which joins will be executed where and performs local joins and other operations accordingly.

By including a query optimizer and an advanced communication system, and through its distributed architecture, Attunity Connect software provides performance that is similar to, and sometimes better than, direct access to a data source.

Attunity Connect software's flexible configuration

Attunity Connect software is designed for optimal performance, whether its components are installed on one machine (client and server) or distributed throughout the network in a multi-tiered architecture. The product consists of client, server, and native driver components, which are available on many platforms and interconnect via an advanced communication system to create a Universal Data Access network.

A driver is required for each type of data source accessed by this software. A server connects data sources to clients and other servers via an advanced communication system. The server manages concurrent connections to data sources and carries out SQL operations (including heterogeneous joins and query optimization). Attunity Connect clients expose a choice of programming interfaces to the applications: ADO/OLE DB, ODBC, and JDBC. Because the software has a peer-to-peer architecture, clients can operate as servers and vice versa.

Attunity Connect components can be deployed to create single or multi-tiered configurations, depending on a customer's business model or application. An example of a single-tier configuration is an application residing on the same system as the data sources. A multi-tiered configuration is an application residing on a client system and data sources residing on a remote server system.

Of course, data sources can reside on the client tier as well. An example of this is an intranet application running on a Windows NT Server-based system with SQL Server that shares data with two remote databases—for example, NonStop SQL database on a NonStop server and RMS database on an OpenVMS platform. With Attunity Connect software's peer-to-peer architecture, a NonStop application can access the RMS files on an OpenVMS platform and the SQL Server database on the Windows NT Server-based system.

conclusion

Today's businesses have the opportunity to leverage a wealth of information, both from inside and outside the organization. But, because of heterogeneous internal computing environments, the plethora of disparate legacy databases and the abundance of diverse external data sources, opportunity has outrun reality. Although a massive amount of valuable information exists, getting to it quickly, simply, and cost-effectively remains a major challenge.

To meet the challenge, HP is providing NonStop system customers with true Universal Data Access across almost any data source and enterprise platform. Attunity Connect software

- Accesses data efficiently wherever it resides, while presenting a single, easy-to-use interface
- Provides access to all data regardless of location, across the Internet and private networks
- Makes up for the syntactic and semantic differences between data sources and compensates for functionality missing in those data sources, providing access to all pertinent data (SQL and non-SQL) using the same semantics

With Attunity Connect software, customers can make their new and legacy databases Internet ready, extend their legacy database investments, utilize existing desktop tools, and deploy new applications that leverage information previously locked up in incompatible internal and external systems.

For more information, go to www.hp.com/go/nonstop.

June 2002, first published 2000. Intel is a registered trademark of Intel Corporation in the U.S. and other countries and is used under license. Microsoft, Windows, and Windows NT are U.S. registered trademarks of Microsoft Corporation. Java is a U.S. trademark of Sun Microsystems, Inc. Oracle is a registered U.S. trademark of Oracle Corporation, Redwood City, California. UNIX is a registered trademark of The Open Group. All other product names mentioned herein may be trademarks of their respective companies. HP shall not be liable for technical or editorial errors or omissions contained herein. The information is subject to change without notice. The warranties for HP products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.

02-0405

©2002 Hewlett-Packard Company

