

# Capturing Your Changed Data With the CONNX Data Synchronization Tool

by Dann Corbit, Senior Software Engineer

October 2004



# Capturing Your Changed Data With the CONNX Data Synchronization Tool

by Dann Corbit, Senior Software Engineer

## Executive Summary

In the modern-day business environment, large organizations or corporations potentially can have extensive numbers of various types of data repositories that operate either in association with off-the-shelf software packages or as custom solutions to business needs. For example, in the supply chain, marketing, and sales areas, an organization might have SAP running on an Adabas database; in the human resources areas, they might have PeopleSoft running on Oracle or Cincom's MRPS program running on VSAM. In addition to this, they might also be running other tools that require permanent data stores from multiple vendors. Naturally, manufacturing information is also needed by the finance department, but the data needs of each of the many departments that comprise an organization vary. The analysis tools or skill sets of the work force for a given department may demand a different database platform for summarization. There may also be ambitious plans to migrate a database system to a different platform, although total corporate demand for data may exceed the bandwidth of the server where the data natively resides. Further complicating the situation is the continuing need for keeping the data safe and secure.

## Fulfilling a Need With Minimal Investment

The CONNX Data Synchronization tool offers a solution to these problems. It is a very simple to use but also highly configurable tool. With the CONNX Data Synchronization tool, security and data integrity are easily maintained and major business problems can be solved with minimal investment.

The CONNX Data Synchronization tool pumps data to a data mart or a data warehouse; users can move a complete snapshot of a pre-defined list of tables. The tool can also incrementally update the current snapshot of data through a simple-to-use scheduling interface. It is also possible to achieve real-time access to your entire organization's data in a virtual data warehouse, due to the functionality that CONNX Solutions' core product, CONNX, brings to the table.

The CONNX Data Synchronization tool is also scalable. It offers multithreaded power that leverages the force of inexpensive modern hardware. Multiple CPU 64-bit machines with thousands of MIPS of throughput can be purchased for a few thousand dollars. Since the CONNX Data Synchronization tool is specifically designed for environments running on multiple machines, the load against production systems can be drastically reduced. If demand warrants, multiple synchronization servers can be added to meet increasing demand. As demand for data expands, inexpensive hardware can be added to whatever degree is required

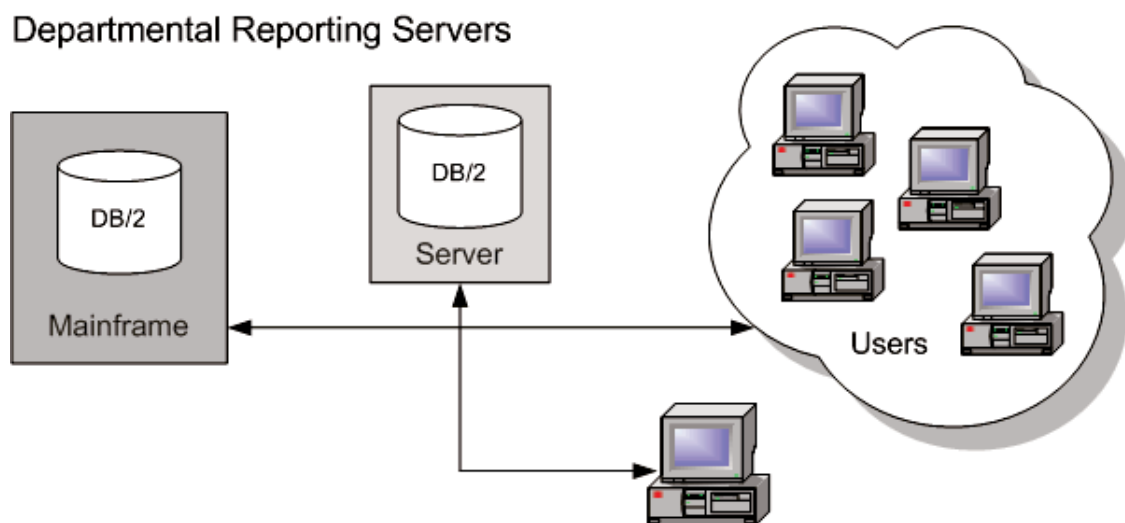
The CONNX Data Synchronization tool is secure. In addition to the security already in place on your database systems, CONNX easily adds whatever sort of additional protection is needed from a central repository in a uniform manner. The tool makes it possible to implement not only table- and column-level security that can be used to restrict access to particular types of information, but even row level security, so that, for instance, an

employee can have access to his own personnel record, but no access whatsoever to the records of other employees.

The CONNX Data Synchronization tool is simple to use, and it has a simple and intuitive graphical interface. It is not unusual to be up and running a production schedule within an hour of installation. To import a source database into our repository is as simple as supplying the connection information. Importing a target database uses the same, simple step. The following step requires that the user simply select the tables to transfer. A mirror image can then be made of the entire database system or any desired subset. Synchronizations can be run at any given time, or on a regular, automated schedule.

Let's review some possible business cases where the CONNX Data Synchronization tool is instrumental in solving business problems.

Figure 1



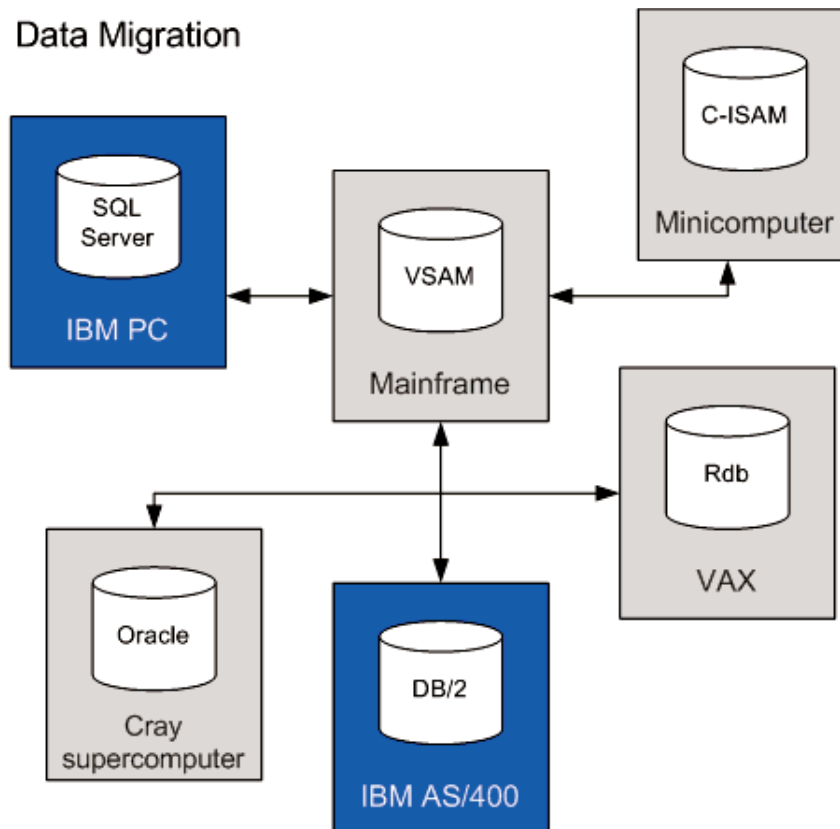
If a user has a need to perform ad-hoc reports which conflict with performance requirements for their mainframe database system, the resolution is to create reporting servers which are automatically synchronized with current data. Queries with large resource requirements or poorly formed queries have no impact on the division's mainframe server database.

## Departmental Reporting Servers

It is not unusual for large organizations to have a need for departmental reporting servers. The usual solution is to write an extract script and a load script that moves data from the production server to form a copy on the target reporting server. However, there are several difficulties with this approach. Moving the entire data volume can be very costly; the batch processes must be monitored for certain conditions, such as running out of temporary disk space or lost connections. The CONNX Data Synchronization tool solves these problems easily with point-and-click selections. Error monitoring is built into the system, which makes it possible to automatically perform different tasks depending upon whether some problem was encountered during the synchronization or whether the usual fully successful synchronization was performed. Additionally, since all data transformations occur automatically, such as changing from big-endian to little-endian format, or from hardware-specific time formats to ANSI/ISO SQL formats, the extra work usually required for such tasks is eliminated.

Primary keys or every key associated with a table can be transferred to a reporting server. It is also possible to allow and preserve customization of the indexes on the target server so that the differing needs of the reporting server users can be efficiently accommodated. The synchronization moves only data that has changed. In a table with millions of rows of thousands of characters each, the difference between moving a few hundred or a few thousand changes is dramatically faster than transferring the entire table. It is easy to separate different groups of tasks into different synchronizations so that each department can get the information they require exactly when they need it - automatically.

Figure 2



Migration of a database system from one operating system and database to another used to take ages to accomplish. Now, with the CONNX Data Synchronization tool set, you can transfer the data and schema from one system to another with tremendous ease. Simply choose the source and target database, and the frequency of synchronization, and it's done. The translations needed to move the schema and data to the new target are performed automatically.

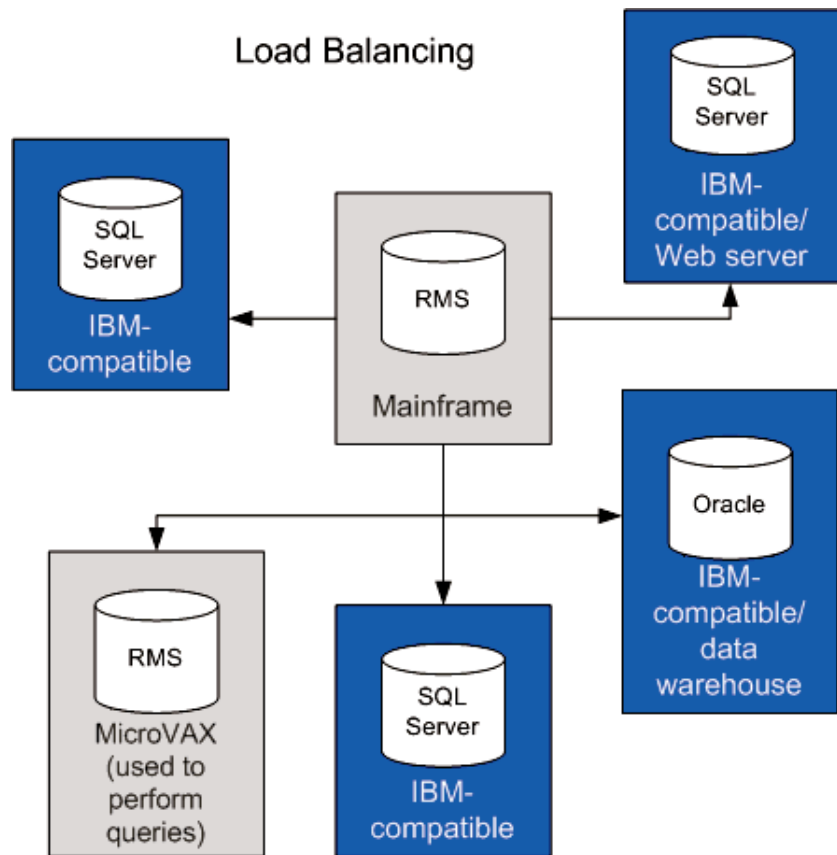
## Data Migration

Most companies face the daunting task of data migration at some point in their product life cycles. As the power of computing resources rises exponentially in accordance with Moore's Law, aging hardware becomes more and more obsolete over time. Users demand intuitive, modern user interfaces rather than the old RS-232 terminal connections common to some older software packages. Market share may be eroded by a clunky and quirky interface written decades ago.

The CONNX Data Synchronization tool helps users solve the issue of data migration. IT staff generally spend endless hours of analysis mapping out all the data transformations required by a typical data migration. To solve this issue, the CONNX Data Synchronization tool can be used to create a target schema that matches the current schema, and then transforms the data types automatically to ANSI/ISO-compliant versions, as well as transfers the data to the target. It can perform all of the data transformations and data conversions without the addition of a single line of code.

Due to the nature of the CONNX tool set, CONNX can be used to develop GUI front ends to your old data stores using Java, VB.NET, or some other front end. Then, when the data has been moved to the new server, the applications remain completely unchanged. All that is required is that the name of the data source be changed. Because the CONNX Data Synchronization tool can automatically synchronize database systems, it is possible to refresh either the source or the target.

Figure 3



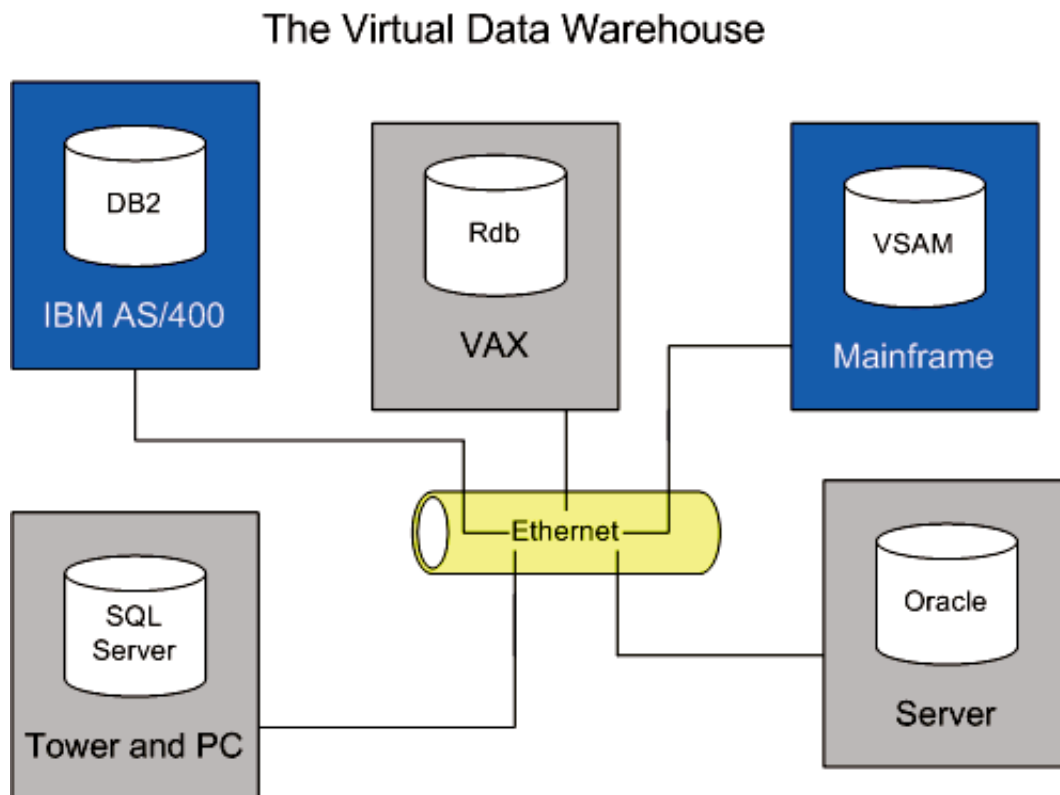
If an Alpha minicomputer running indexed RMS files is the host for business applications, the load can be balanced for those who need access to the data by synchronization of the RMS files to several other database systems. By reduction of the load on the main server, users can retain high performance and ready access to the needed information.

## Load Balancing

Another conspicuous need of corporate data is load balancing. A given hardware platform may be able to efficiently handle 1,500 users, but there may be 10,000 users requiring connectivity. One way to work around this problem is to mirror the data to multiple systems and give end users access to the target servers. In this way, we can smoothly scale up to rising demand by simply adding more mirrors. The database type of the target system can be any type with which users are comfortable. The target database choice is completely flexible and choosing one database over another has virtually no impact at all. All conversions of data to appropriate representations are handled with no effort by the tool user. This feature grants increased flexibility to decision makers who can postpone the need to perform a data migration for several years by drastically reducing the load on the host mainframe server.

Since only the changes to the data are propagated to the target servers, the network bandwidth is not impacted in the way it would be if full loads were performed each time. Because a smaller volume of data is transferred, there is also a very large time savings. Less CPU usage can be a large cost savings when a mainframe is involved and especially when the time is purchased or expensed by the resource volume consumed by a department.

Figure 4



There are times when real time access to data is required. Instead of trying to move all the corporation's data simultaneously to some single repository, CONNX can be used as a virtual data warehouse. CONNX can give continuous real time access to all data stores within an organization

## CONNX as a Virtual Data Warehouse or EII

Another use for CONNX is the creation of a virtual data warehouse or Enterprise Information Integration (EII) tool. With CONNX, every database used within the organization can be imported into a central repository. The resulting CONNX Data Dictionary (CDD) gives live, real-time access to every data source without movement of any data. Because using CONNX results in all tables being viewed as peers, it is not necessary to write programs to join data from one database to another. A SQL query can transparently join data from one database to another as though they were within a single database system. The ability to create a virtual data warehouse is inherent in the core product of CONNX and comes gratis with the CONNX Data Synchronization tool.

Ideal for executive decision-making, reports can be run across all data collections corporate-wide. The database systems do not even have to reside in the same state or even the same country.

## Conclusion

The CONNX Data Synchronization tool is cost-effective software that enhances any organization's ability to solve business problems resulting from database issues. Its ease of use, coupled with its elevated performance and security features, make the CONNX Data Synchronization tool and CONNX the ideal data integration and data management solution. The software package can be used to speed up or stave off migration. It is also known for its adherence to open database standards, including .NET, OLE DB, JDBC, ODBC, and is completely platform-neutral.

## Summary of Product Functions

The CONNX Data Synchronization tool provides several unique product functions.

- Extracts data on demand  
Completes data extraction as snapshots from selected files.
- Full reloads on demand  
Purges all prior snapshot data in the data warehouse, replacing it in the data warehouse with a completely new snapshot.
- Incremental updates on demand  
Detects which records have been added, deleted, or changed, and inserts, updates, or deletes the corresponding data in the data warehouse snapshot tables at a fraction of the usual time. Initial full update from source to target may take 2 to 3 times longer than the equivalent transfer using FTP (3.5 hours) because of the following beneficial costs:
  - ◆ Insert of data into SQL Server
  - ◆ Creation of CRCs
  - ◆ Storage on CONNX Store
  - ◆ Transformation of the data types from source schema to target schema.
  - ◆ Other costs, such as location on a Wide Area Network

However, the incremental update usually runs up to 10 times faster than the full update, a significant decrease in the time it takes to synchronize.

- Performs a scheduled incremental update  
Detects which records have been added, deleted, or changed, and inserts, updates, or deletes the corresponding data in the data warehouse snapshot tables. Each file in the synchronization process may be scheduled independently, or as a group.
- Performs a scheduled full reload  
Purges all prior snapshot data in the data warehouse. A completely new snapshot of the data is placed in the data warehouse. Each file in the synchronization process may be scheduled independently, or as a group.
- Creates datamart model schema  
Automatically performs a complete analysis of all source database schemas and performs complete, applicable transformations for the target datamart schema.
- Creates datamart model transforms  
Automatically creates datatype transformations of all data elements from source schema to target schema.

## About CONNX Solutions, Inc.

CONNX provides businesses secure read/write real-time access to all enterprise data from any platform as if all the data existed in one relational database. This technology has been referred to as a “federated database”, or a “virtual database”. All data is then accessible using standard ANSI SQL and any standards-based application. CONNX acts as a reusable data access framework for projects throughout the enterprise. CONNX supports Adabas (currently OS/390, z/OS, and Windows platforms); Microsoft SQL Server (on all platforms); Sybase (on all platforms); Informix (on all platforms); C-ISAM, DISAM, and Micro Focus (on Linux, AIX, Sun, Solaris, and SCO operating systems); IBM VSAM (on z/OS, OS/390, MVS, and VSE); IBM DB2 (on all platforms); Oracle (on all platforms); RMS, Rdb, and Codasyl DBMS (on OpenVMS and VMS platforms); PostgreSQL (on all platforms); Dataflex; POWERflex; and any OLE DB, ODBC, or JDBC data source. For more information, visit [www.connx.com](http://www.connx.com) or e-mail [info@connx.com](mailto:info@connx.com)