# Université IBM i 2018

## 16 et 17 mai

IBM Client Center Paris

## S33 – Réplication, échange et transformation de données DB2 en temps réel

Aurélie GODEC & Patrick DESRUELLE

**SYNCSORT**

*Aurelie.godec@syncsort.com* & *Patrick.desruelle@syncsort.com*

# Plan de la présentation

Partage des données en temps réel (Migrations, BI,…)

- MIMIX Share

- Quick-SmartData

Anonymisation des données

- Quick-Anonymizer

- Nota : Si vous souhaitez obtenir la version PPT avec ses commentaires en français, n'hésitez pas à nous adresser un email accompagné de vos remarques éventuelles (Merci)

-  Je confirme que MIMIX Share permet également de générer automatiquement la base cible.

# MIMIX Share & Quick-SmartData :
## *Easy, Real-Time Data Sharing Between Databases*

**Aurélie GODEC**

Senior Director, Product Management, IBM i

Aurelie.Godec@syncsort.com

**Patrick DESRUELLE**    (+33 6 0739 6437)

Regional Sales Director  France-Benelux-Afrique FR

Patrick.desruelle@syncsort.com

# Syncsort, Advancing Data
## We organize data everywhere, to keep the world working

### Data infrastructure optimization

## Syncsort Optimize

MFX® for z/OS
ZPSaver Suite
EZ-DB2
EZ-IDMS
DMX & DMX-h
DMX AppMod
DL/2
Zen Suite
athene®
athene SaaS®

### Data integration & quality

## Syncsort Integrate

Ironstream®
Ironstream® Transaction Tracing
DMX & DMX-h
DMX Change Data Capture
Trillium Software System
Trillium Quality for Big Data
Trillium Precise
Trillium Cloud
Trillium Global Locator

### Data availability & security

## Syncsort Assure

MIMIX Availability & DR
MIMIX Move
MIMIX Share
Quick-EDD/HA
iTera Availability
Enforcive, Cilasoft, CSi
Quick-SmartData
Quick-Anonymizer

### NPD, R&D, NV

## Syncsort Advance

New Ventures

# Today's Topics

- Common data sharing needs & practices

- Traditional data sharing methods

- MIMIX Share overview

- MIMIX Share at work

- Partnering with Syncsort

# Trends and Practices in Data Sharing
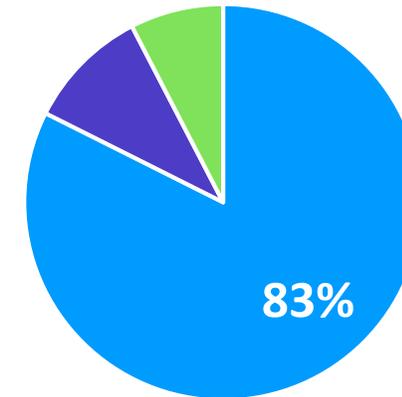
# Businesses Have Multiple Databases

## Multiple databases are the norm

- Merger or acquisition
- Choice of multiple apps or databases for best of breed solutions
- Combination of legacy and new databases
- Multi-organization supply chain

## IT infrastructures are heterogeneous

- Database platforms
- Operating systems
- Hardware

**Does your organization rely on multiple databases?**

**83%**

■ Yes   ■ No   ■ I don't know.

**Does your organization share data between multiple databases?**

**73%** of those with multiple databases share data among them

Source: Vision Solutions ' 2017 State of Resilience Report

# Varied Business and IT Goals
## for Data Sharing

- Protecting performance of production database by offloading data to a reporting system for queries, reports, business intelligence or analytics

- Offloading data for maintenance, backup, or testing on a secondary system without production impact

- Consolidating data into centralized databases, data marts or data warehouses for decision making or business processing

- Maintaining synchronization between siloed databases or branch offices

- Feeding segmented data to customer or partner applications

- Migrating data to new databases

- Replatforming databases to new database or operating system platforms

**For what business purpose does your organization share data between databases?**
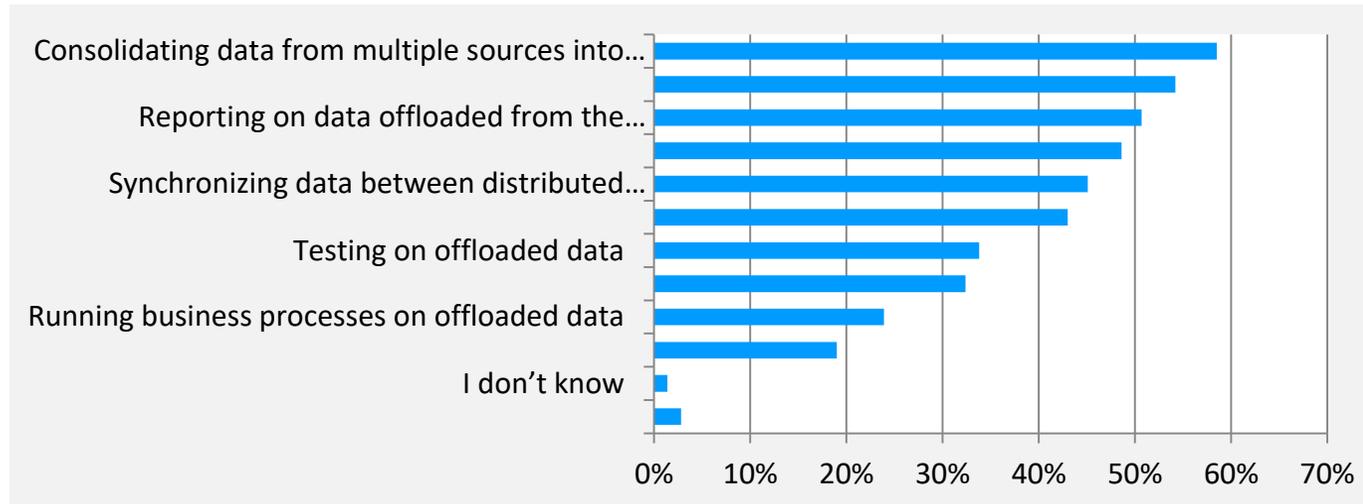
Source: Vision Solutions ' 2017 State of Resilience Report

# How Does You Need to Move Data ?

- Enabling queries, reports, business intelligence or analytics without production impact ?

- Populating centralized databases, data marts or data warehouses ?

- Feeding real-time data to employees, customers or partners ?

- Keeping data from siloed databases in sync ?

- Reducing the impact of database maintenance, backup or testing ?

- Replatforming to new database or operating systems ?

- Consolidating databases ?

# Traditional
# Data Sharing Methods

# Traditional Methods for Sharing Data

## Direct network access

- Reporting on production servers across the network during business hours

    *Issue :* Negatively impacts network and database
    performance – resulting in user complaints!

## Off-hours reports and extractions

- Run reports off-hours or perform nightly ETL processes to move data to a reporting server

    *Issue :* Business operates on aging data until next extraction

    *Issue :* Difficult to find acceptable time to perform an extraction

## ETL (Extract-Transform-Load) Processes

- FTP/SCP/file transfer processes
- Manual scripts
- Backup/restore
- In-house tools

    *Issue :* Periodic, not real-time, delivery of data

    *Issue :* Labor intensive to create processes and tools

    *Issue :* Expensive to develop and maintain

    *Issue :* Prone to errors

# In-House ETL Scripts and Processes Are Not Free

- **Upfront development costs**
  - Development of code to perform database extraction, transformation, and load
  - Additional requirements for additional pairings, schemas, etc.

- **Test system expenses**
  - Hardware and storage resources
  - Database licenses for test systems
  - Add-on products, e.g. gateways

- **Maintenance costs**
  - Ongoing enhancements for altered schemas, additional platforms
  - Testing new database and OS releases
  - Cross training and documentation to reduce turnover risk

- **Lost opportunity costs for other initiatives**

# Introducing
# MIMIX Share

# MIMIX Share Transforms Your Business

**Breaks down barriers between databases**

- Same or different database management systems
- Same or different operating systems
- Physical, virtual or cloud platforms
- Across any distance

**Makes data sharing easy**

- Replicates database changes in real time
- Transforms and enhances data during replication
- Supports leading database and operating systems
- Offers a variety of replication architectures
- Easy graphical UI – no programming required!

**Quickly returns your investment**

- Stronger decision making
- Greater business productivity
- Ability to choose more cost-effective infrastructure
- Frees IT to focus on other business initiatives
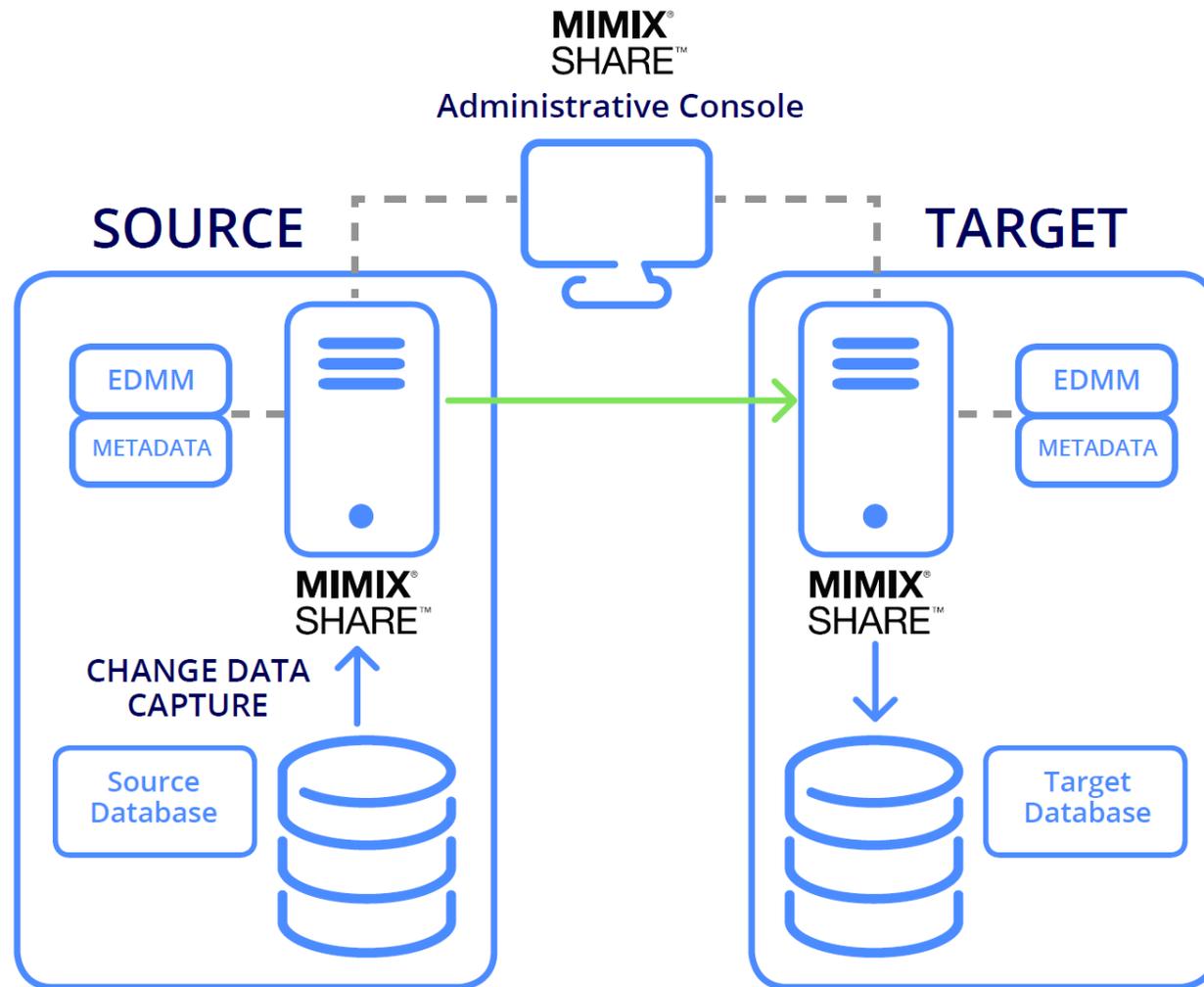
**MIMIX SHARE**
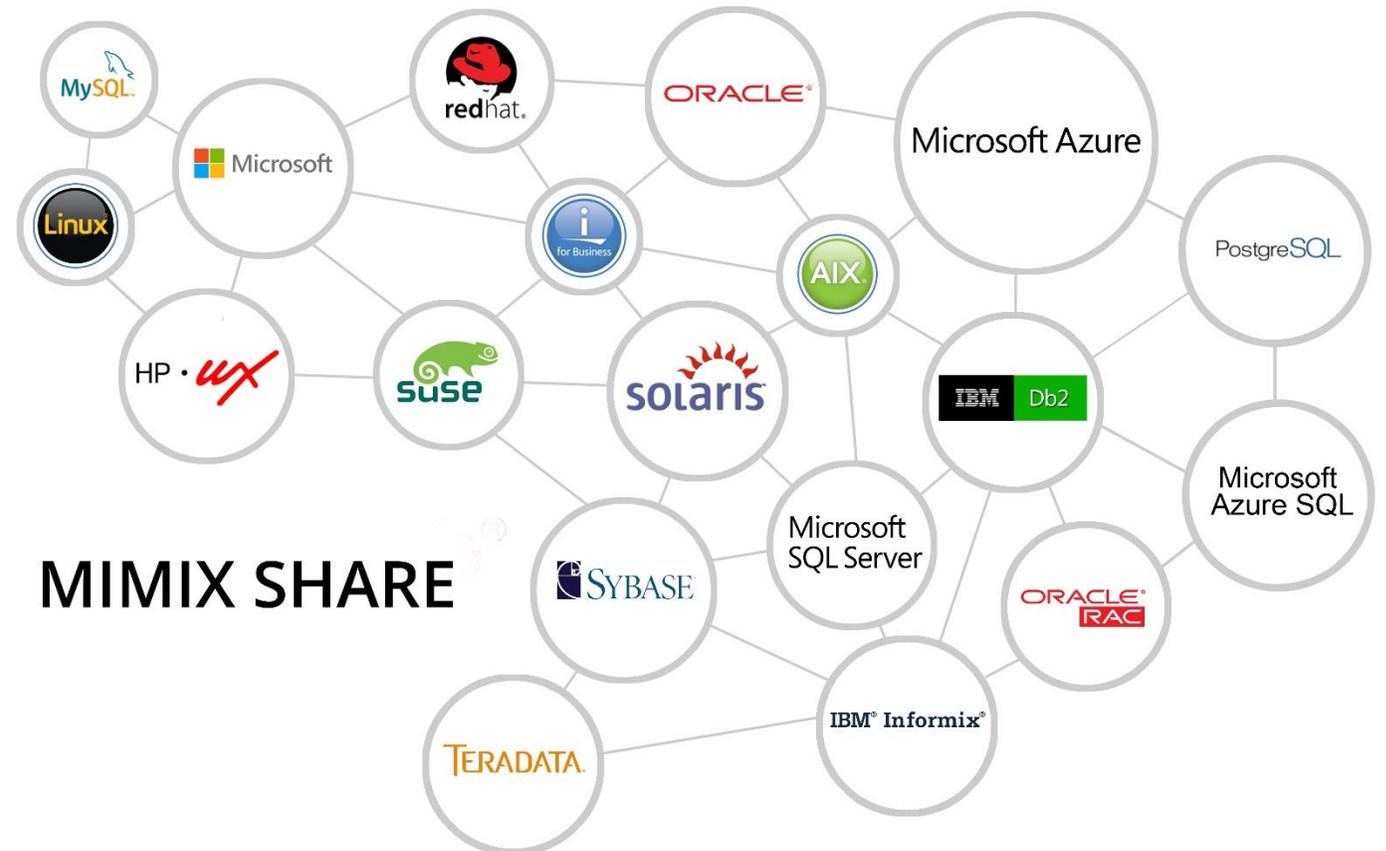
# Replaces Manual Processes

- Point & click graphical user interface
- Single view of data across databases and operating systems
- Simple, model-based configuration
- 80+ pre-built, click-and-go data transformations
- Transformations can be added through Java-like scripting
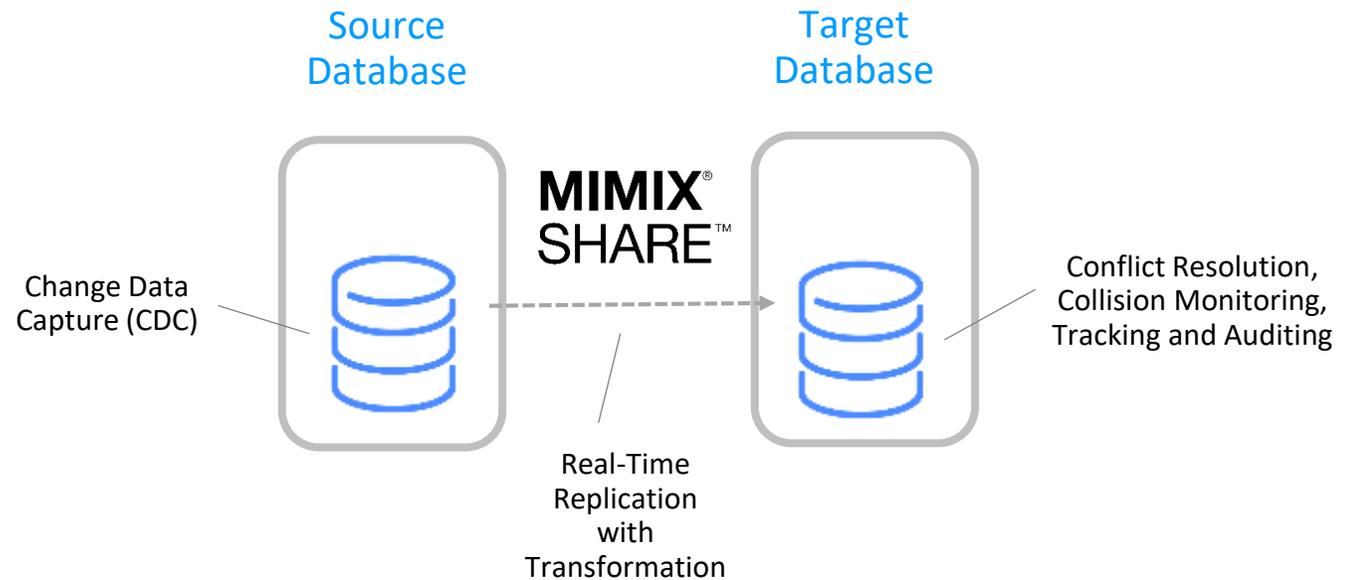- No programming required
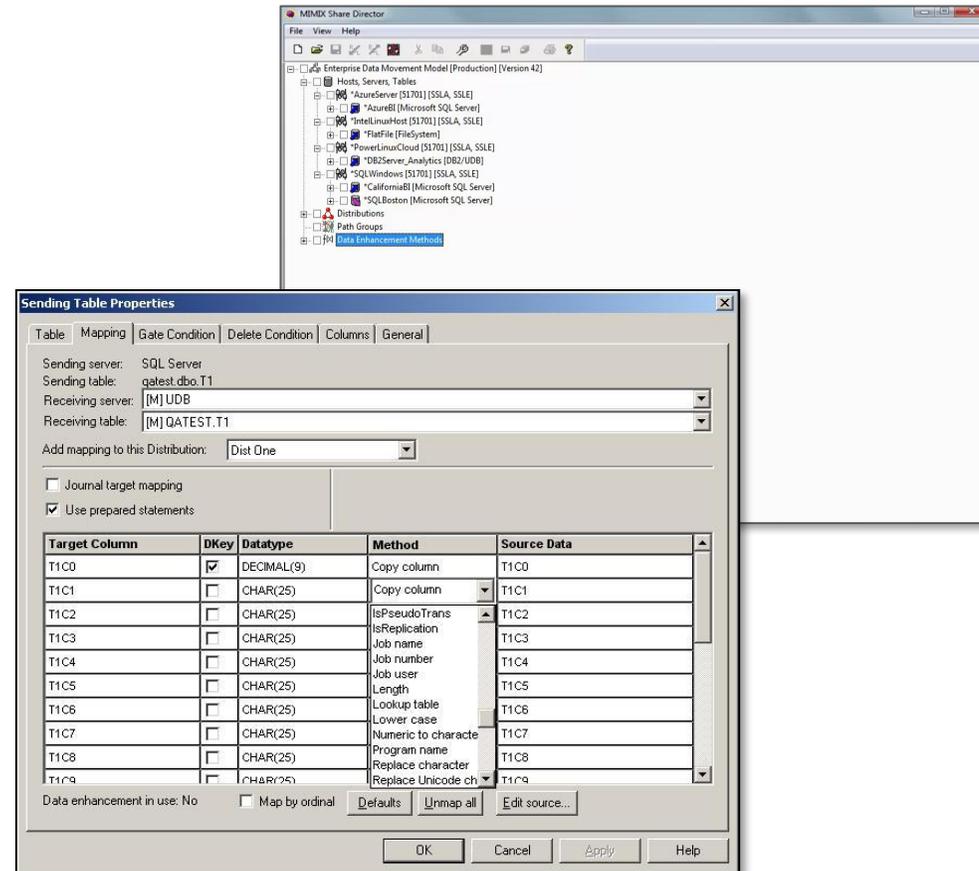
# High-Level Architecture

# Replicates Data in Real Time

- Change Data Capture (CDC) captures database changes immediately and quickly replicates them to another database(s)

- Only changed data is replicated to minimize bandwidth usage

- Automatically extracts, transforms and loads data into target database without manual intervention or scripting

- Ensures write order consistency and guaranteed delivery

- Ensures data integrity with conflict resolution and collision monitoring

- Enables tracking and auditing of transactions for compliance

Source Database

Target Database

MIMIX® SHARE™

Change Data Capture (CDC)

Conflict Resolution, Collision Monitoring, Tracking and Auditing

Real-Time Replication with Transformation

# Technical Overview
# MIMIX Share

# Change Data Capture (CDC)

Data is captured and shared as it changes, immediately and quickly replicated to other databases as specified
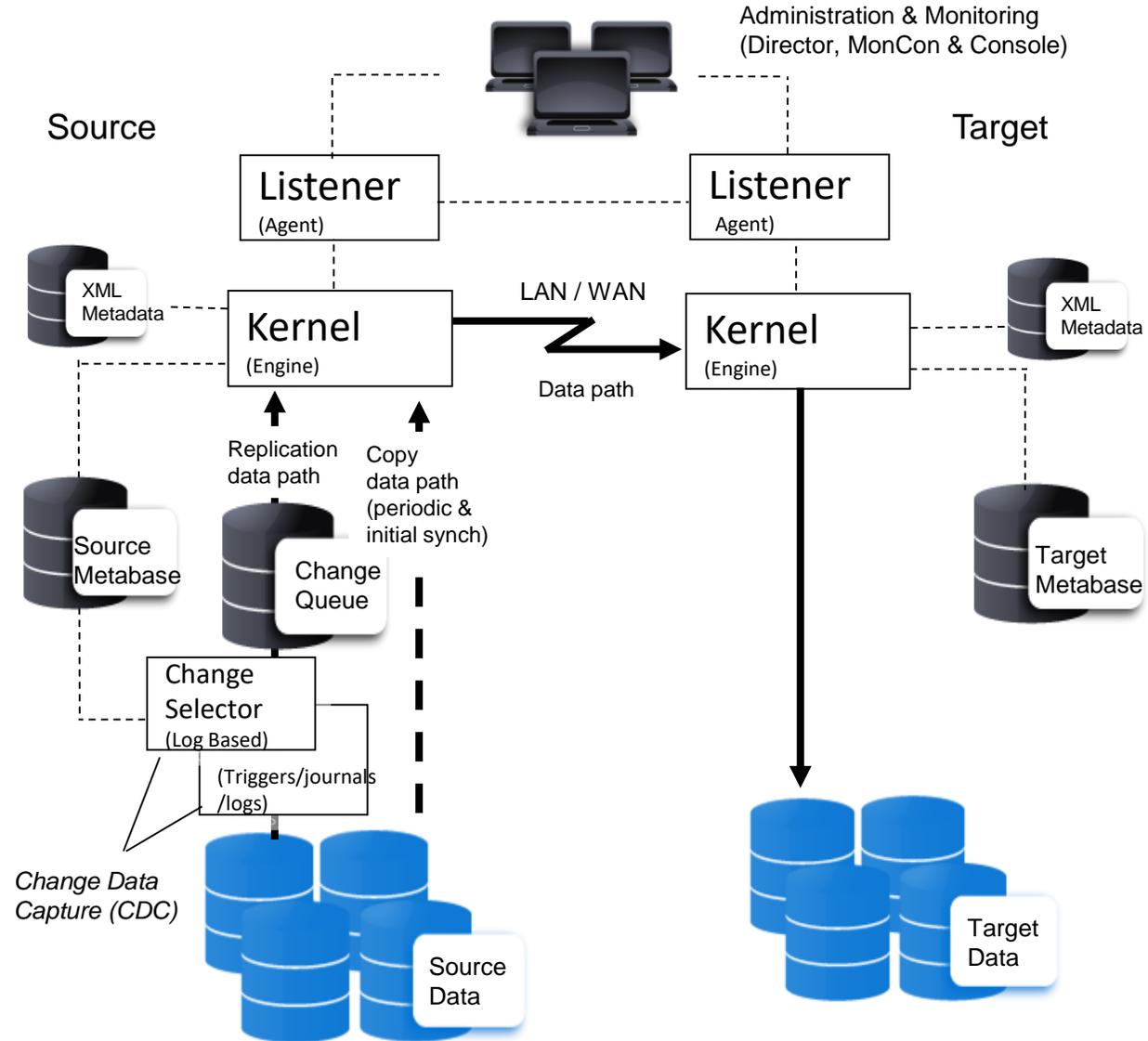
Only changed data is transferred, so efficient

Automatically handles the extraction, transformation (if any) and load to target database (no intervention or manual scripting)
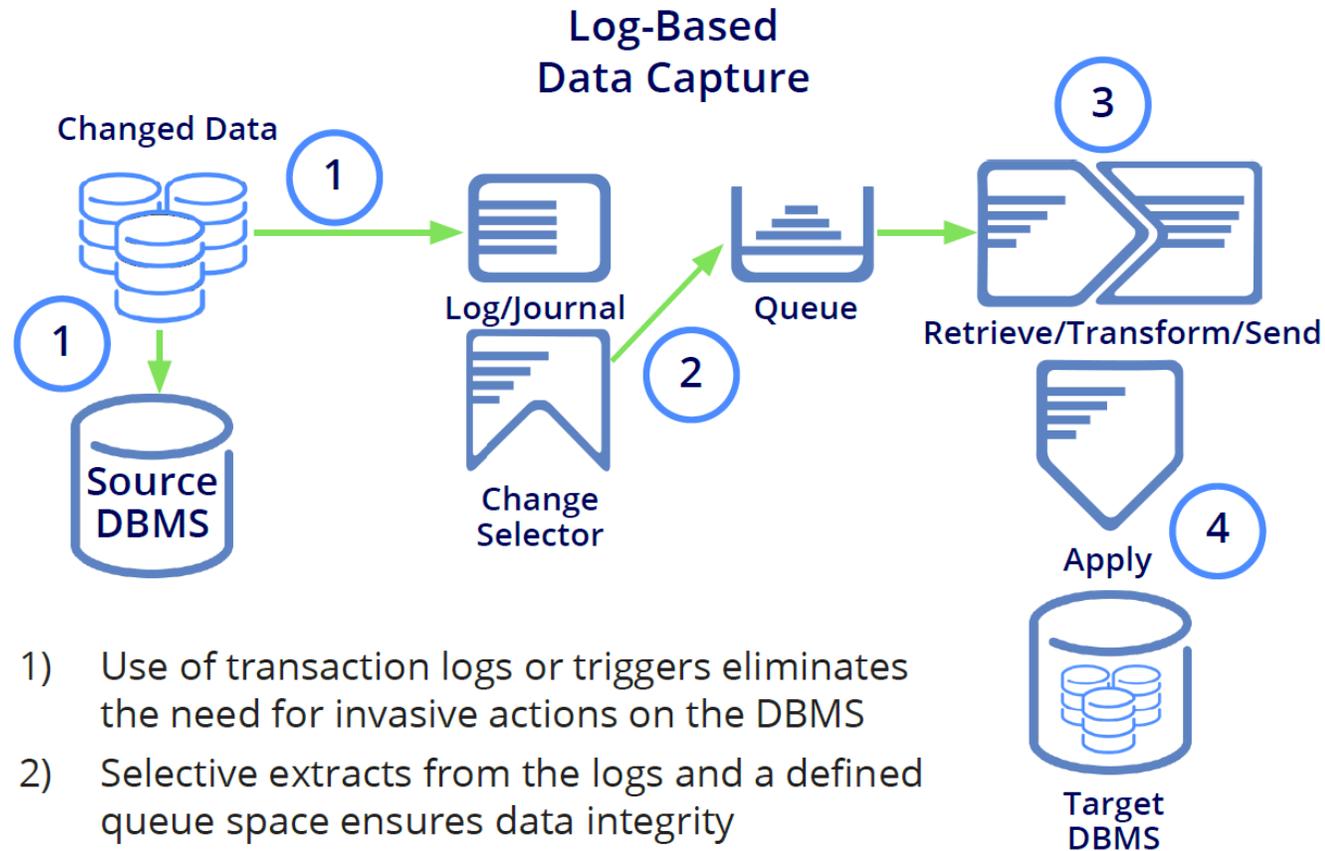
Ensures write order consistency and guaranteed delivery

# How Does Change Data Capture Work?

# Log-Based Data Capture

**Log-Based Data Capture**

Changed Data



Source DBMS

Log/Journal

Change Selector

Queue

Retrieve/Transform/Send

Apply

Target DBMS

1) Use of transaction logs or triggers eliminates the need for invasive actions on the DBMS
2) Selective extracts from the logs and a defined queue space ensures data integrity
3) Transformation in many cases can be done off box to reduce impact to production
4) The apply process returns acknowledgment to queue to complete pseudo two-phase commit

## Did you know?

MIMIX Share can leverage published Log or Journal standards to identify and capture the change before copying to the MIMIX Share Queue.

The MIMIX Share Queue ensures that data integrity is maintained and zero data loss occurs in the event of a dropped connection during file transmission.

# MIMIX Share Components & Monitoring



**EDMM**
All data replication rules including data mappings and data manipulations are stored internally using a scripting language in an Enterprise Data Movement Model.



**MonCon**
The monitoring and control interface is one of several ways to monitor the status of replication in real time. MonCon lets you do this locally through a graphical monitor. There is also WebMonCon that allows monitoring through a web browser.

# Guarantees Information Accuracy

## Ensures ongoing integrity

- Changes collected in queue on source
- Moved to target only after committed on source
- Ensures write-order-consistency retained
- Queues retained until successfully applied
- No database table locking
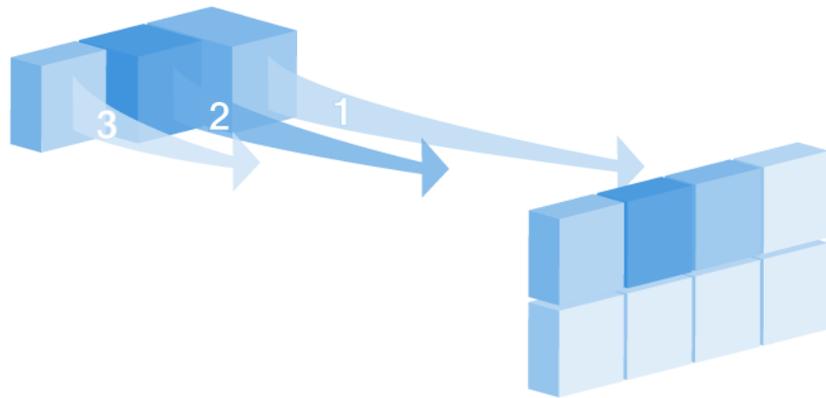
## Ensures failure integrity

- Automatically detects communications errors
- Automatically recovers the connection and processes
- Alerts administrator
- No data is lost

**SMTP Alerting**
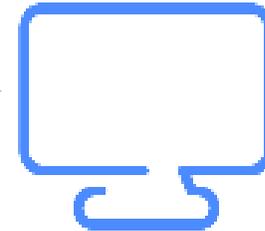
# Accurate Tracking & Data Auditing

Detects and resolves conflicts
- Maintains data integrity

Model verification
- Validates date movement model
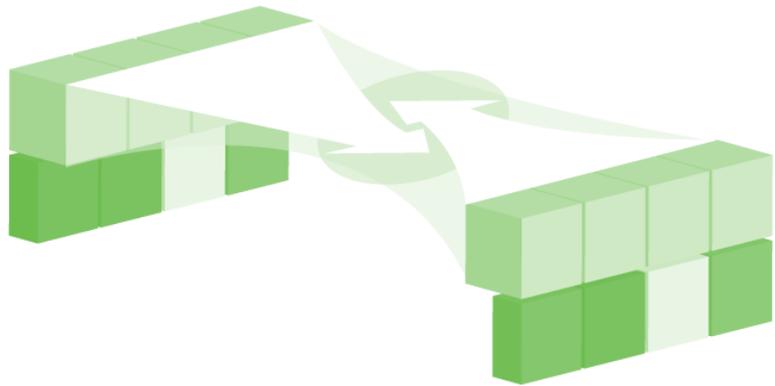
Audit Journal Mapping tracks all updates and changes
- Records
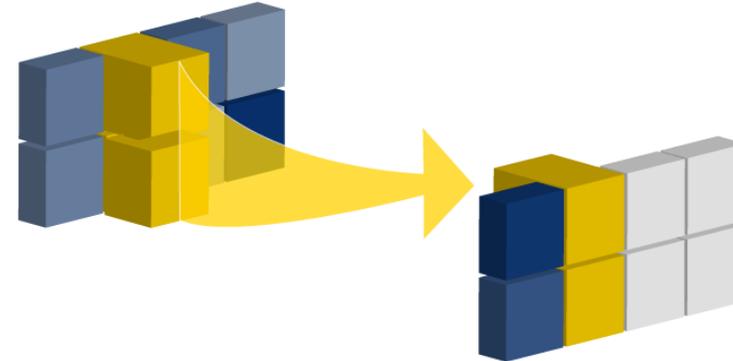    - Before and after values for every column
    - Type of transaction
    - Type of sending DBMS
    - Table name
    - User name
    - Transaction information
- Records to flat file or to database table
- Can assist with SOX, HIPAA audit requirements

# Lets You Share Exactly WHAT You Need

Filters determine what data gets moved

- Select specific column and table

- Select specific rows and table
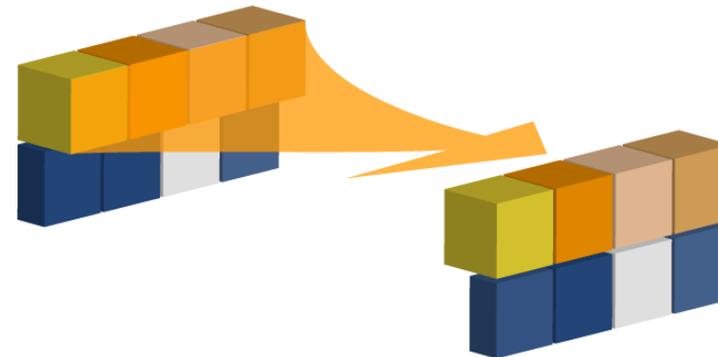
# Transforms Data Exactly HOW You Need To

- Transforms data into useful information

- 80+ built-in transformation methods

- Field transformations, such as:

  - DECIMAL(5,2)

  - nulltostring(ZIP_CODE,'00000')

- Table transformation, such as:

  - Column merging

  - Column splitting

  - Creating derived columns

- Custom lookup tables

- Create custom data transformations using powerful Java scripting interface

# Maps Columns

**Source Host**

connected to TCP/IP network

Database Server
Db2

**Target Host**

connected to TCP/IP network

Database Server
MS SQL Server

table mapping

**CUSTPF**

**customer_master**
(SQL table)

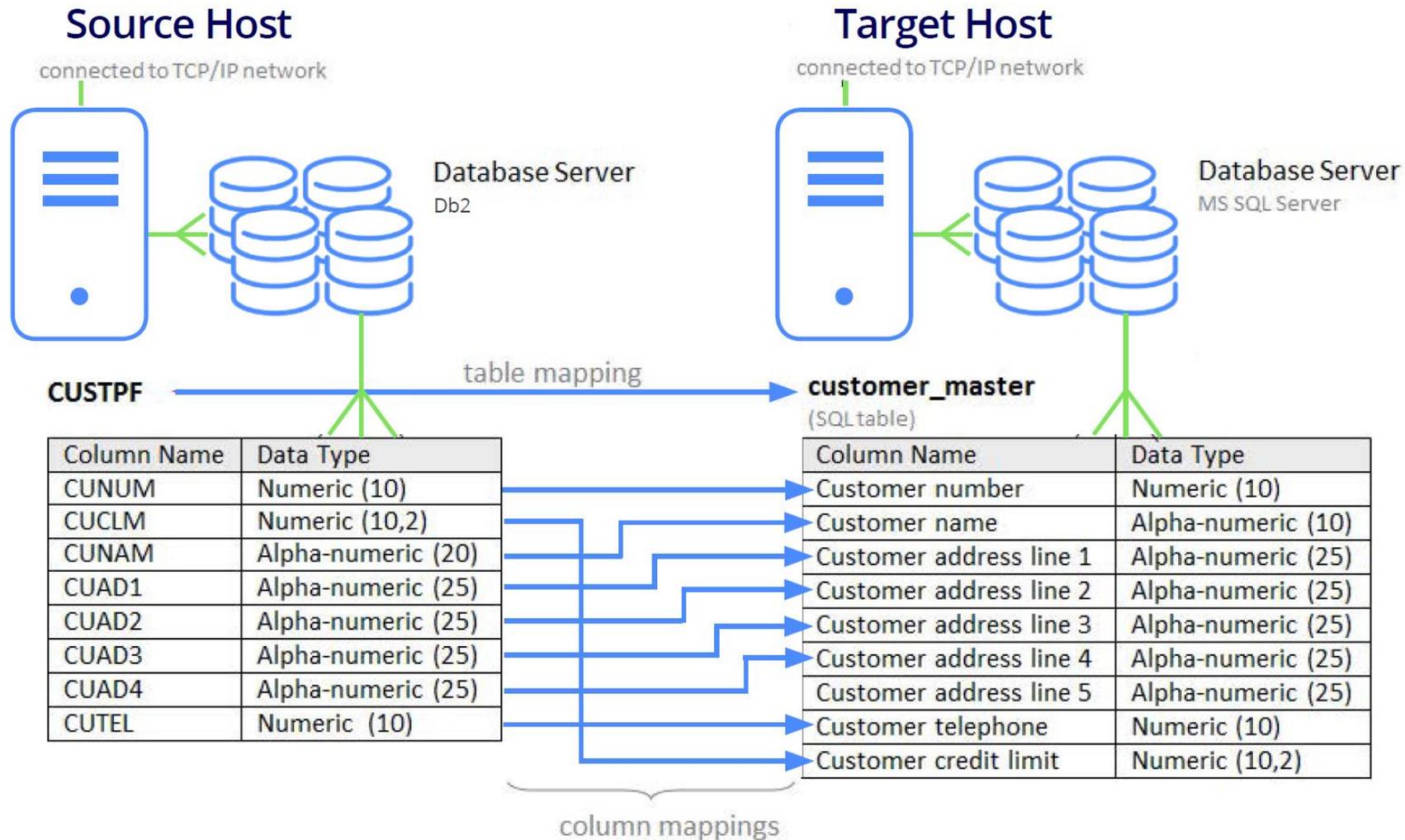| Column Name | Data Type |
|---|---|
| CUNUM | Numeric (10) |
| CUCLM | Numeric (10,2) |
| CUNAM | Alpha-numeric (20) |
| CUAD1 | Alpha-numeric (25) |
| CUAD2 | Alpha-numeric (25) |
| CUAD3 | Alpha-numeric (25) |
| CUAD4 | Alpha-numeric (25) |
| CUTEL | Numeric  (10) |

| Column Name | Data Type |
|---|---|
| Customer number | Numeric (10) |
| Customer name | Alpha-numeric (10) |
| Customer address line 1 | Alpha-numeric (25) |
| Customer address line 2 | Alpha-numeric (25) |
| Customer address line 3 | Alpha-numeric (25) |
| Customer address line 4 | Alpha-numeric (25) |
| Customer address line 5 | Alpha-numeric (25) |
| Customer telephone | Numeric (10) |
| Customer credit limit | Numeric (10,2) |

column mappings

# Manipulating Data



**Target Column Expressions**
Fast way to specify simple and medium complexity data manipulations
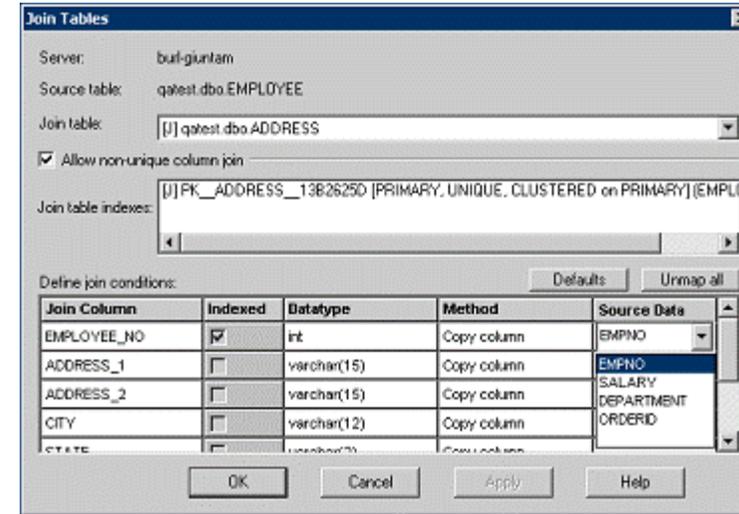


**Join Tables**
It's possible to join tables together and build a new table on the target.
Eg. Join the invoice header table with the customer master table to get address columns

```
-- mappings.rsl
CREATE TABLE MAPPING
TARGET DVCOLE.EMP AT AXQ4
FROM SOURCE astest.dvams.claims90 AT I10HP4
EMPID [DKEY]: claim_no, EMPNAME: claimee, STATE: location
TRAN_ID: "Copy input" (TRANSACTION_ID),
TRAN_SEQ: "Copy input" (TRANSACTION_ROW_SEQUENCE)
;
CREATE TABLE MAPPING
TARGET astest.dvams.claims90 AT I10HP4
FROM SOURCE DVCOLE.EMP AT AXQ4
claim_no [DKEY]: EMPID, claimee: EMPNAME, location: STATE
;
CREATE TABLE MAPPING
TARGET SAIXD1.qadba.emp AT SAIXD1
FROM SOURCE DVCOLE.EMP AT AXQ4
empid [DKEY]: EMPID, empname: EMPNAME, state: STATE
;
CREATE TABLE MAPPING
TARGET SAIXD1.qadba.emp AT SAIXD1
FROM SOURCE DVCOLE.EMP AT AXQ4
empid [DKEY]: EMPID, empname: EMPNAME, sal: Foo(EMPID, EMPNAME)
;
CREATE TABLE MAPPING
TARGET SAIXD1.qadba.emp AT SAIXD1
FROM SOURCE astest.dvams.claims90 AT I10HP4
empid [DKEY]: claim_no, empname: claimee, state: location
;
CREATE TABLE MAPPING
TARGET astest:dvams.claims90 AT I10HP4
FROM SOURCE SAIXD1.qadba.emp AT SAIXD1
claim_no [DKEY]: empid, claimee: empname, location: state
;
```

**Replication Scripting Language (RSL)**
All data manipulations are stored in a easy to read scripting language, which can be exported, edited and imported for advanced manipulations by power users. Also a big time saver in sophisticated enterprise environments with lots of tables.

# Supported Platforms for MIMIX Share

## Leading Operating Systems

- IBM i
- IBM AIX
- HP-UX
- Solaris
- IBM Linux on Power
- Linux SUSE Enterprise
- Linux Red Hat Enterprise
- Microsoft Windows, including Microsoft Azure

## Leading Databases

- IBM DB2 for i
- IBM DB2 for LUW
- IBM Informix
- Oracle
- Oracle RAC
- MySQL*
- Microsoft SQL Server
- Microsoft Azure SQL*
- PostgreSQL*
- Teradata*
- Sybase

*\* Target database only*

# Syncsort's Solutions

**MIMIX Share**

**Quick-SmartData**

Combined strengths deliver more for your business!

# Supported Platforms for MIMIX Share & Q-SmartData

## Leading Operating Systems

- IBM i
- IBM AIX
- HP-UX
- Solaris
- IBM Linux on Power
- Linux SUSE Enterprise
- Linux Red Hat Enterprise
- Linux openSUSE
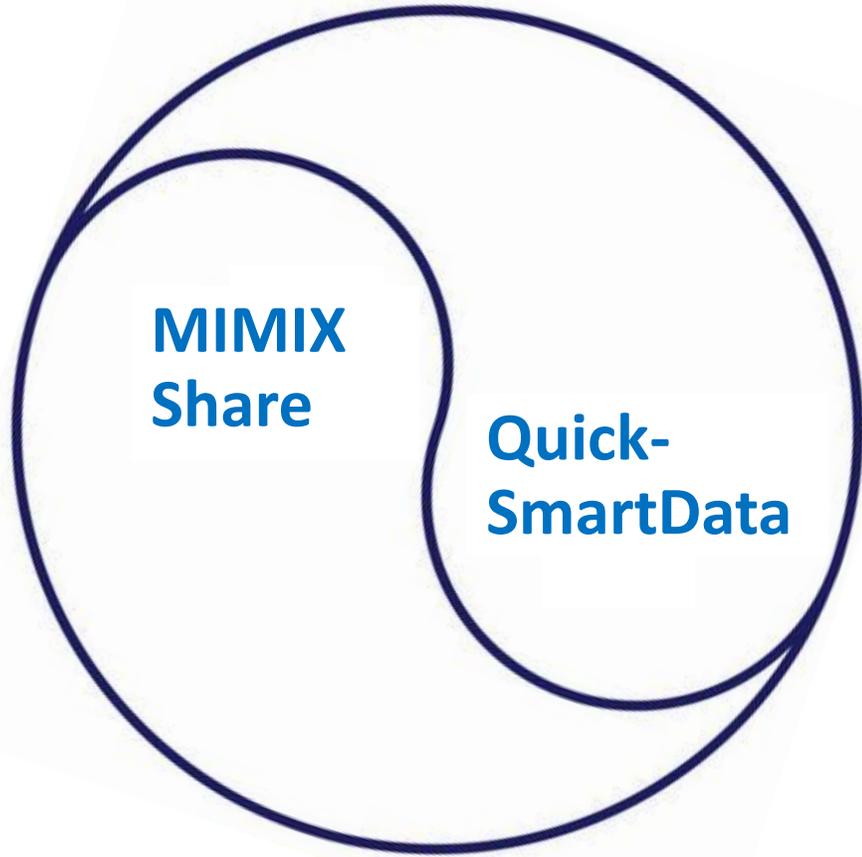- Linux Fedora
- Microsoft Windows, including Microsoft Azure
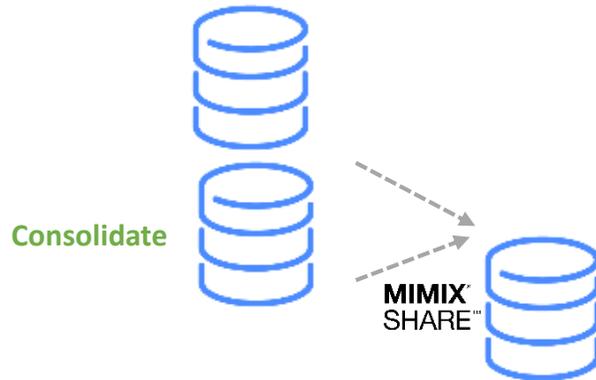
## Leading Databases

- IBM DB2 for i
- IBM DB2 for LUW
- IBM Informix
- Oracle
- Oracle RAC
- MySQL
- Microsoft SQL Server
- Microsoft Azure SQL*
- PostgreSQL
- Teradata*
- Sybase
- MariaDB
- Netezza*

*Target database only*

# Flexible Replication Options



One Way

Distribute
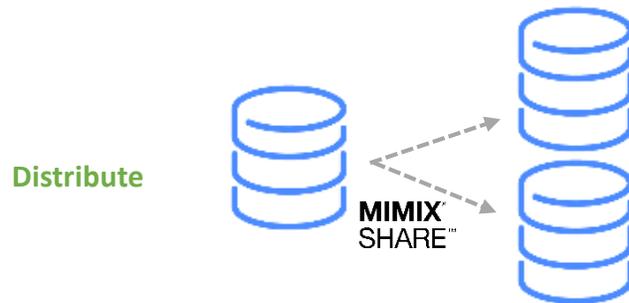
Consolidate

Choose a topology or combine them to meet your data sharing needs

Two Way

Bi-Directional

Cascade
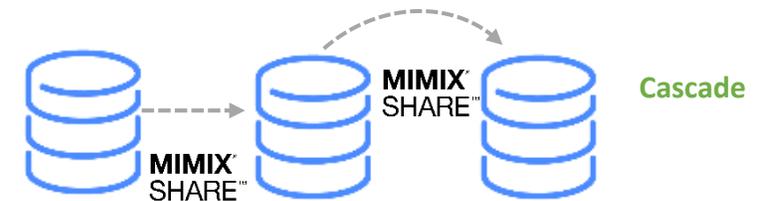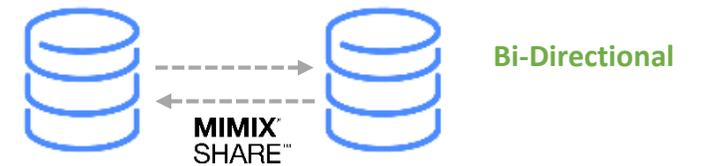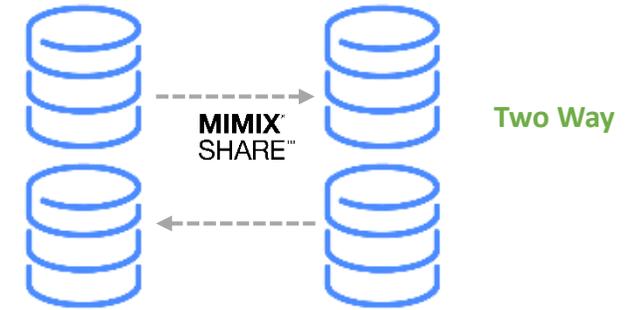
# MIMIX Share Addresses All Your Data Sharing Needs

✓ Facilitates real-time queries, reporting and business intelligence

✓ Enables centralization or consolidation of data

✓ Transforms data for smooth data flow between databases

✓ Keeps distributed applications and data in sync

✓ Feeds real time data to mission critical applications

✓ Offloads data for maintenance, testing and backup

✓ Migrates legacy data to new platforms

✓ And more !

# MIMIX Share
# in Action

# Offload Reporting from Production Use Case

Many cost effective tools available
on MS SQL server
platform for query reports

Production System

Real time CDC
replication
with transformation

Offload Query System

**MIMIX SHARE**

IBM System i DB2
Lawson M3 (Movex)

MS SQL Server

Query reports

Data Warehouse load

Reduce CPU and I/O overhead
on production system to
improve user response times

Data is already partially
'scrubbed' and available
for loading data warehouses
and data marts without
performance impact on
production system

RED WING
SHOES
EST. 1905

# Centralized Reporting Use Case



Casino 1
IBM System i
DB2

Casino 2
IBM System i
DB2

Casino 3
IBM System i
DB2

Casino 4
IBM System i
DB2

Casino 5
IBM System i
DB2

Casino 6
IBM System i
DB2

**MIMIX®**
**SHARE™**

Real time CDC replication
with transformation

- Customer loyalty
- Amounts paid
- Amounts won
- Time at the table
- Time at the machine

**Single Data Warehouse Database**
Windows Cluster
MS SQL Server

Business intelligence

# Database Migration Use Case

Old System

Real time CDC
replication
with transformation

New System

MIMIX®
SHARE™

IBM i
DB2 for i
JDE (standard)

IBM i
DB2 for i
JDE (Unicode)
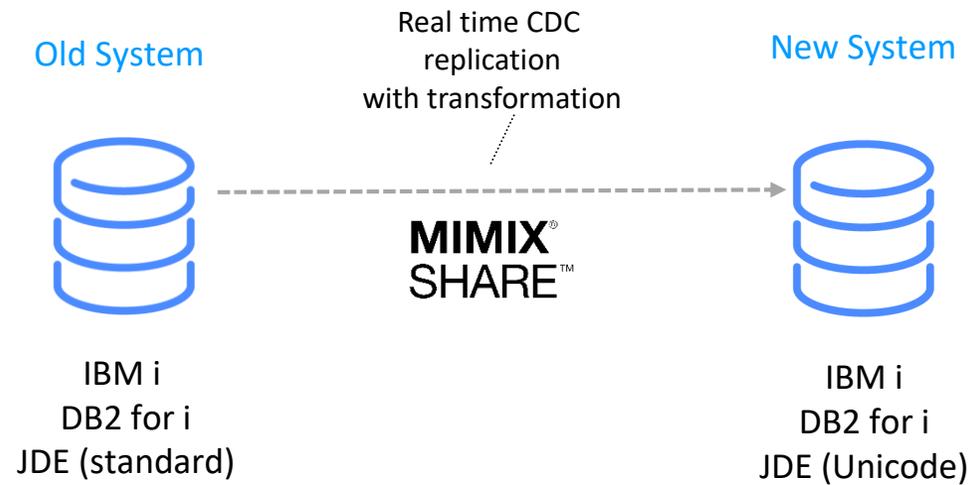
**Manufacturing
Company**

# Database Replatforming Use Case

Transformation between different OS and database platforms

Users are moved to new server in phases over a period of time
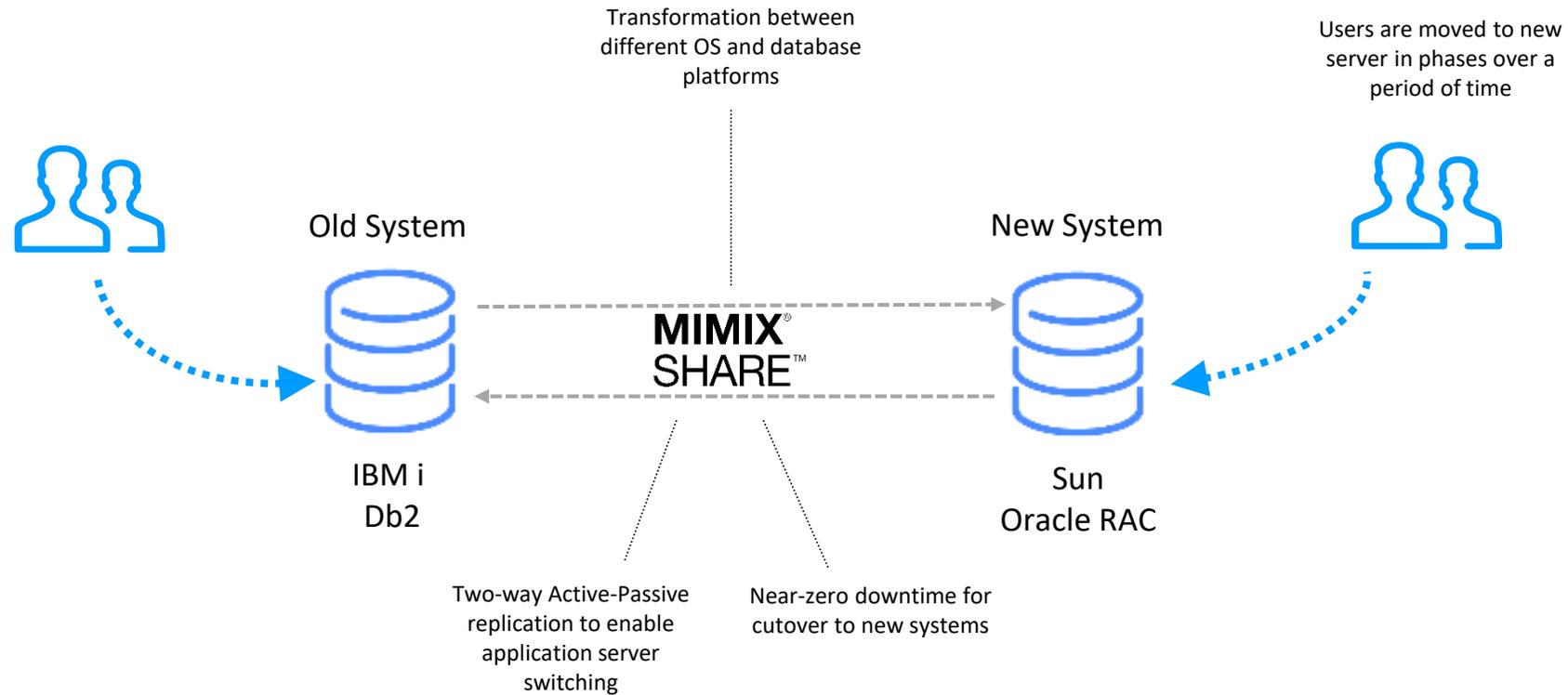
**Large Insurance Company**

Old System

New System

**MIMIX® SHARE™**

IBM i
Db2

Sun
Oracle RAC

Two-way Active-Passive replication to enable application server switching

Near-zero downtime for cutover to new systems

# Gradual Database Replatforming Use Case

Transformation between different OS and database platforms with **completely different schemas**

100s of partners moved to new server after training nd at their own pace

America II Corp

Old System

New System

**MIMIX® SHARE™**

IBM i
DB2

Windows
SQL Server

Active-Active replication eliminated need for hard cutover and enabled partners to move back and forth between systems

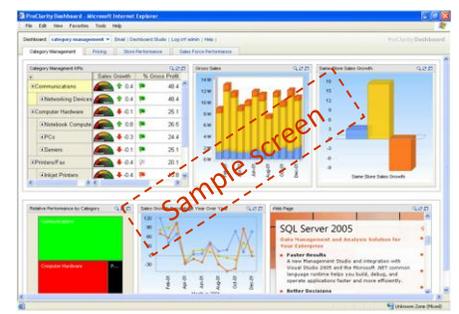True zero downtime for migration to new systems

# Reporting in Azure Use Case

**1** End users enter data into their local applications.

**2** MIMIX Share captures, transforms, and replicates database changes in real-time.

**3** Replicated data is merged into a single, consolidated MS SQL reporting database in the Azure cloud.

**Source**

**MIMIX SHARE™**

**Target**

*Sample screen*

**4** Business Intelligence apps and reporting tools access the consolidated information from the Azure-based database.
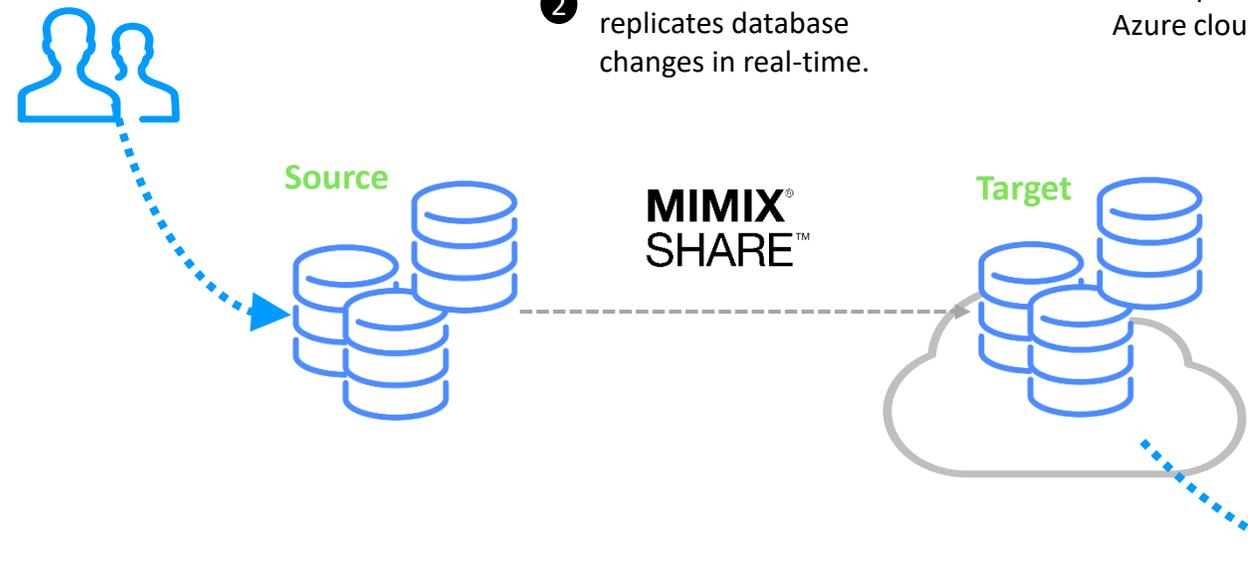
# Application Integration Use Case

**Online Banking**

**Microsoft SQL Server**

**2** MIMIX Share replicates the transactions in real time to Db2/400

**IBM System i Db2**

**1** Customers enter new banking transactions on line. They get captured in SQL Server

**MIMIX**® **SHARE**™
Bi-directional replication

**5** Customers view processed transactions on-line

**3** A back-office batch application processes incoming transactions and updates data.

**4** MIMIX Share replicates the processed transactions back to SQL Server in real time.

# Additional Use Cases



ERP SYSTEM

eCOMMERCE &
WEB PORTALS

Customer Orders
Payment Details

**MIMIX**® SHARE™

Product Catalogue
Price List

**MIMIX**® SHARE™

DR /
BACKUP

**MIMIX**® SHARE™

DATA EXCHANGE
WITH OUTSIDE VENDOR
(FLAT FILE)

Outside
Vendor

**MIMIX**® SHARE™

TEST & AUDIT
ENVIRONMENT

# Customer Success Stories

"The time required to replicate data between applications has dropped dramatically. Some data used to take a full day to replicate. Now it's available immediately."

*Harsh Anand, Manager of Groupware and IT Support Services*

*Amway*

"We can't live without MIMIX Share. When it comes to setting up the replication model, there's a lot of metadata to help you with that and its quite intuitive. If we had to go back to the old way of doing things it wouldn't be a pleasant situation."

*Andy F., Senior Systems Analyst*

*Getinge North America*

# Customer Success Stories

"MIMIX Share helps keep our students and their families informed in real-time and enables our teachers and administrators to perform faster, better data analysis. Real-time access to select information provides us with analytic capabilities comparable to those at Fortune 500 companies."

*Phyllis Chasser, Ph.D., Senior Data Warehouse*

*Analyst, Broward County Public Schools*

"MIMIX Share gives us many advantages as a company, allowing us to make decisions quickly with shared, real-time data when we need to."

*Xavier Majem, IT Manager*

*Grupo Uriach*

# Quick-Anonymizer

# Why Quick-Anonymizer?

- **Supports GDPR compliance** - No personally identifiable information (PII) outside the production context

- **Data consistency** - Key notion is respected

- **Control of information flow** – Feeds data in real time or on a schedule to other systems for
  - Testing
  - Development
  - Education
  - External systems for ERP editors

- **Strong technical foundation** – Powerful replication engine

46

# How Quick-Anonymizer Works

- Anonymization rules are applied on production data **before** they're sent to target system

- No real data gets out of the production system – **ever**!

- Field level process

- **Irreversible**

- Benefits from all replication features
  - Simple, fast, real-time replication
  - Synchronization and verification

# Quick-Anonymizer Technologies

**Processes all the events in sequence**

Journal BD

AUDIT (QAUD
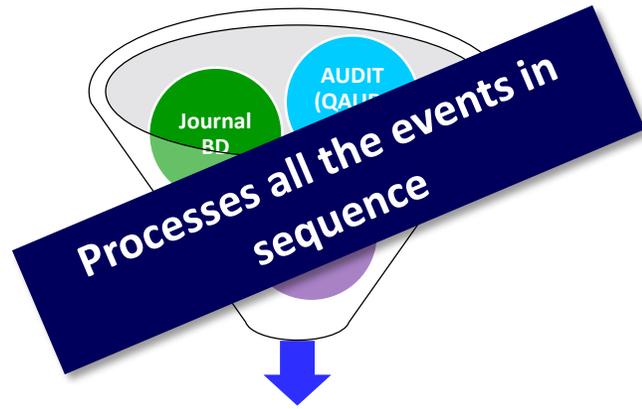
**Sent to the target IN SEQUENCE using one single virtual journal**

Send

Apply

Acknowledge

**Configuration and supervision from the source system**

**Real-time SYNCHRONOUS process without inconvenience**

**Synchronization with no locks. Verification through audits/repairs.**

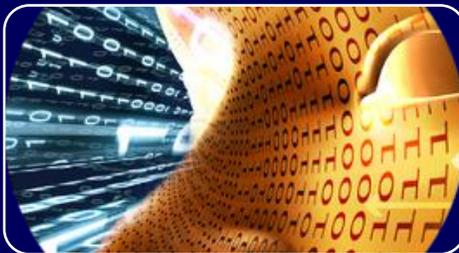**REPAIR**

# Anonymization Options

## SCRAMBLE

- Partial:  21st John Doe Street  ->  56fd John Doe Street
- Full:    21st John Doe Street  ->   56fd Furh Sop Rjekdt

## MASKING

- Partial:  +1 448 924 364  ->   +1 XXX XXX XXX
- Full:    +1 448 924 364  ->  ** *** *** ***

## PCI/LUHN

- Credit Card:                 5141 4268 8381 9357        ->      6153 3472 7303 8395
- Social Security Number:    1 99 11 75 342 125 34       ->      2 85 04 84 345 124 32
- Banking Account Number:   30001007941234567890185  ->  62044200262666688884470
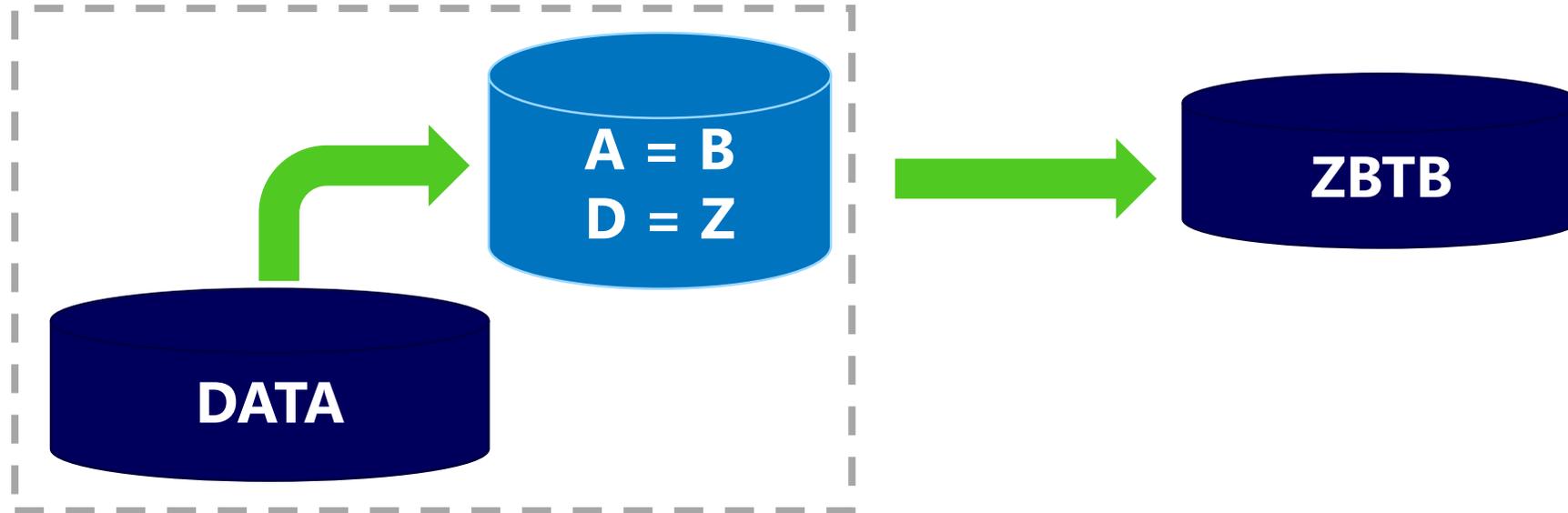
# Anonymization Options

## EXIT PROGRAM

- Custom transformation rules
- Referential database is required in this case

# Fields Finder

**Reminder of your selection :**

| Libraries | Files | Names | SQL Leng... | Descriptions | Long Field Names | Field headings |
|-----------|-------|-------|-------------|--------------|------------------|----------------|
| DPH | DHWRTE | RTTE02 | 7 | RESUL TEST 02 | RTTE02 | RESUL TEST 02 |
| DPH | DHWRTE | RTTE04 | 7 | RESUL TEST 04 | RTTE04 | RESUL TEST 04 |
| DPH | DHWRTE | RTTE06 | 7 | RESUL TEST 06 | RTTE06 | RESUL TEST 06 |
| DPH | DHWRTE | RTTE08 | 7 | RESUL TEST 08 | RTTE08 | RESUL TEST 08 |

4 rows

**①  How do you want to anonymize the selection**

| Operation : | Type : | Subs / Char : | Start : | Length : |
|-------------|--------|---------------|---------|----------|
| SCRAMBLE | STANDARD | | 0 | 0 |
| *Informations about the different operations* | *Informations about the different types* | *Character substitution.* | *First character position.* | *Number of character substitute.* |

**②  Where do you want to anonymize the selection**

Environment :

**③  Save your anonymization**

← Back          💾 Save anonymization

51

# How is Quick-Anonymizer Unique?

- The only real-time solution that is native to the IBM i

- Easy and fast to install, configure and manage

52

# Professional Services
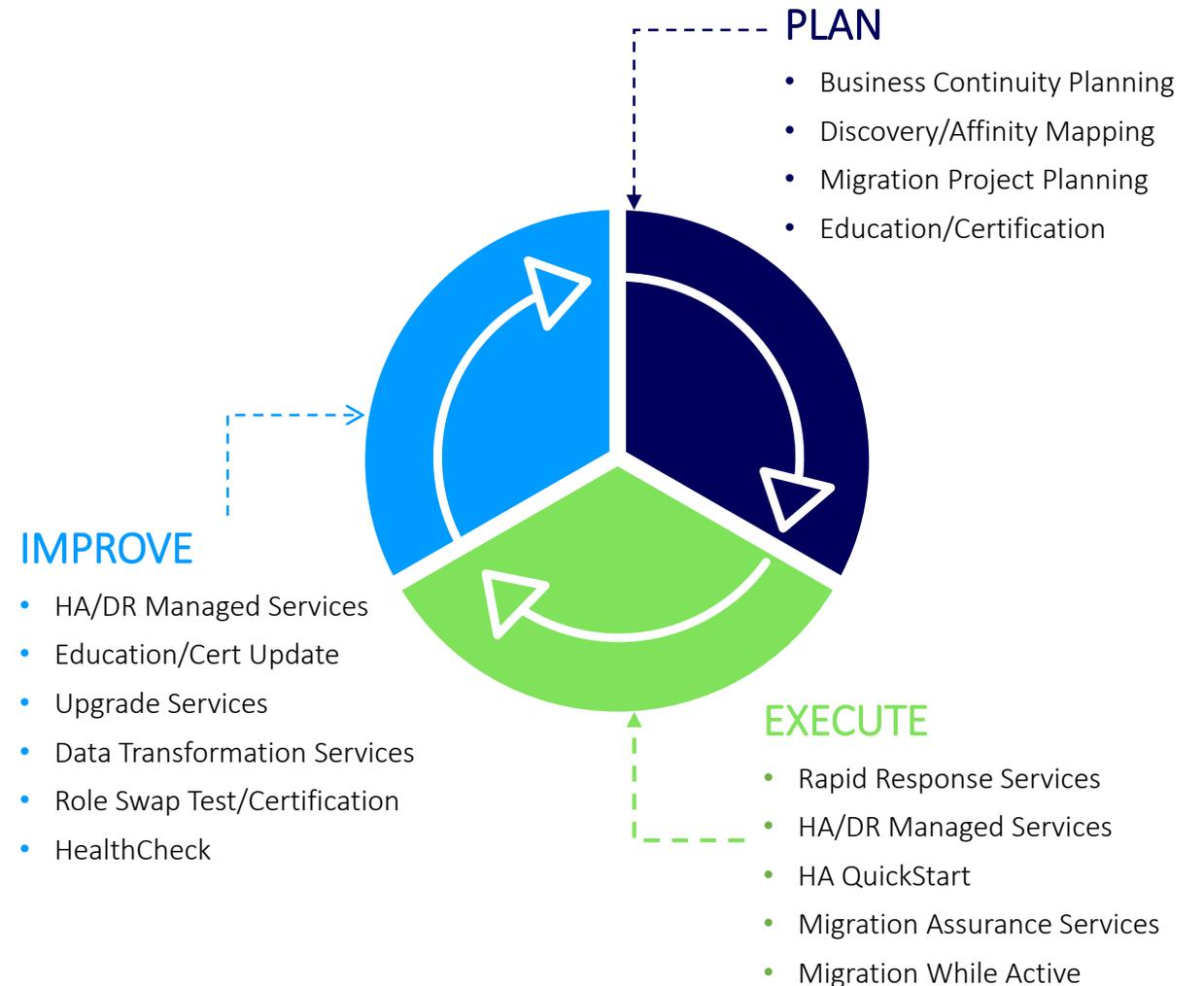
## MIMIX Share Quick Start

- Expert implementation tailored to customer needs
- Solution planning workshop
- Installation and configuration
- Skills transfer

## Project-based Custom MIMIX Share Services

- Migration
- Replatforming
- Business process integration
- and more

## MIMIX Share Update

- Expertly customized upgrade from a previous release
- Includes knowledge transfer
- Cost varies by environment

### PLAN
- Business Continuity Planning
- Discovery/Affinity Mapping
- Migration Project Planning
- Education/Certification

### IMPROVE
- HA/DR Managed Services
- Education/Cert Update
- Upgrade Services
- Data Transformation Services
- Role Swap Test/Certification
- HealthCheck

### EXECUTE
- Rapid Response Services
- HA/DR Managed Services
- HA QuickStart
- Migration Assurance Services
- Migration While Active

# Education

## MIMIX Share Administration Class

- Three day onsite or remote, instructor led education

- Main topics
  - Concepts
  - Architecture
  - Installation and Planning
  - Director
  - Monitoring and Control
  - Replication queues
  - Director Advanced
  - Director Data Enhancement
  - Post Installation

- eLearning – MIMIX Share Administration
  - Online, self-paced
  - Covers same topics as classroom version