

# 실시간 분석과 의사 결정을 위한 인포매티카 데이터 복제 솔루션 소개

2012. 4. 24

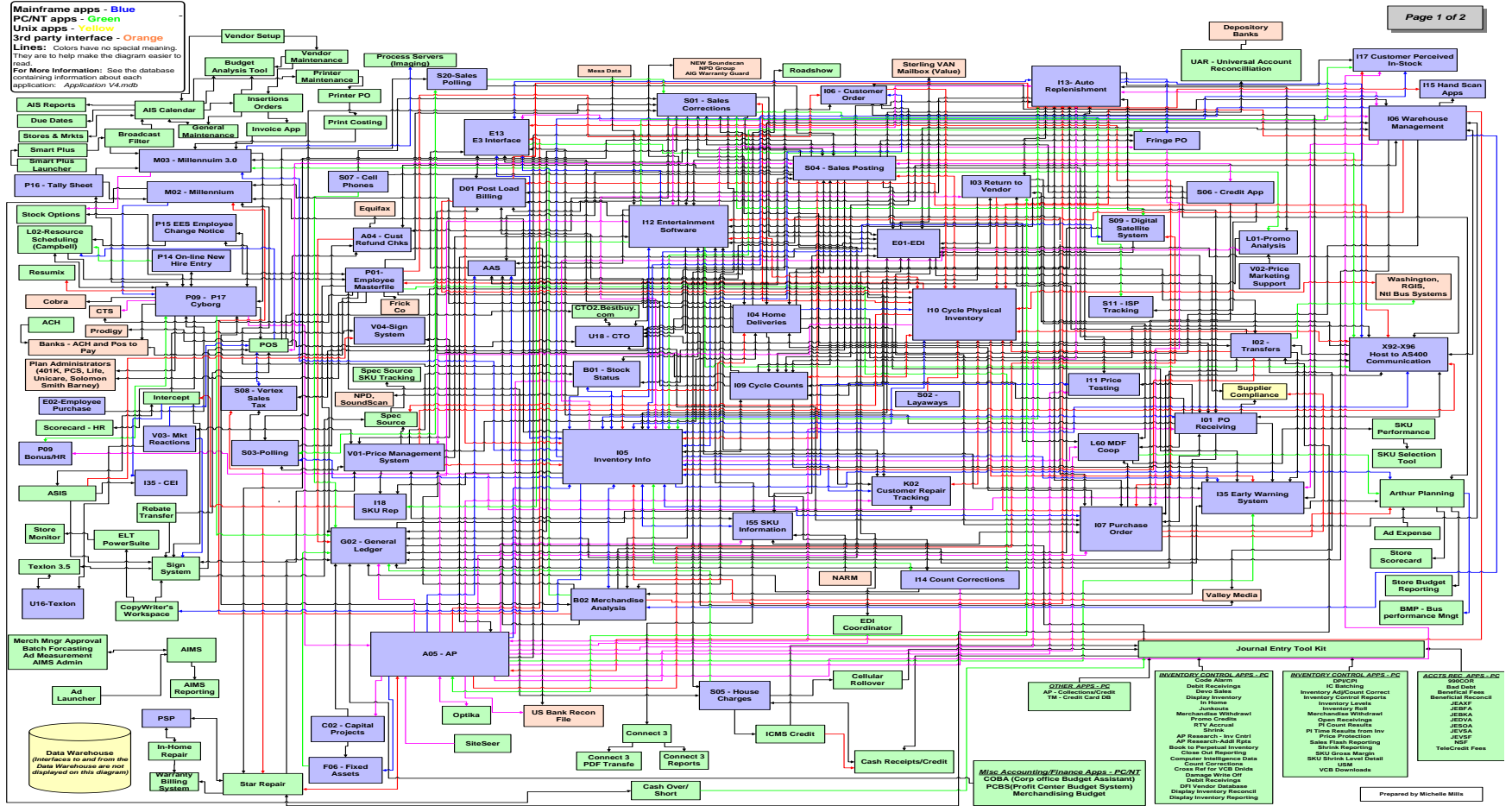
한국인포매티카

임정혜

# Data Integration

## Customer Challenges

# The Data Integration Challenge



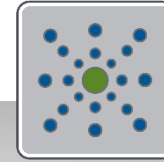
# Business Value through Trustworthy, Actionable, Authoritative Information Assets



Data Quality



Complex Event Processing



Master Data Management

## Information Infrastructure



Cloud Data Integration

NETEZZA	PeopleSoft	flockdb	MapR	HPA	HIPAA
SYBASE	Microsoft	CONVEA	Data	SWIFT	
Greenplum	SAP	mongoDB	EML	IBM	ACORD
TERADATA	SIEBEL	Cassandra	PDF		
Hyperion	ORACLE	Noxaj		SEPA	
VERTICA	IBM	Impatica		NACHA	

B2B Data Exchange

Ultra Messaging

Enterprise Data Integration

ILM

## Data Infrastructure

# Informatica Universal Data Replication

*The Most Comprehensive Solution in the Market Today*

## Informatica Data Replication

Broad, multipurpose heterogeneous transactional replication solution -  
*Informatica's strategic enterprise-class replication product offering*

## Informatica Fast Clone

Rapid extraction and loading of Oracle data into heterogeneous targets including  
broadest set of warehouse appliances

*Informatica's snapshot replication for unlocking Oracle data*

## PowerExchange CDC

Change Data Capture for ETL and Data Quality for incremental update requiring complex  
transformation

*Informatica's Change Capture Technology for Trickle Feeding*

## Cloud Data Replication

Replication of Salesforce data between Cloud and On-Premise

*Informatica's Cloud based Replication for SaaS*

# Business Drivers for Data Replication

## The Needs and Pains

### Real-Time Data for Decision Making

Fresh data always available for reporting  
Analyze current and historical information using an Active Data Warehouse

### Reduce IT Spend

Choose cost effective systems for analysis  
Offload reporting to lower cost systems

### Continuous Availability

Eliminate downtime from hardware/software upgrades  
Eliminate downtime from application maintenance



# What Market is saying

## Customers and Analysts agree...

At the CAB (Customer Advisory Board) meetings, the **# 1** asked for enhancement to Informatica's portfolio has been Data Replication

“Data replication is a flexible technology that can be used for many purposes, including data integration, data movement for data warehouses and ODS, reporting server, migrations, and high-availability requirements. Transactional log-based replication is one of the most common usage scenarios”

*Forrester TechRadar™ “Q1 2010 Enterprise Data Integration report”*

“Data Replication tools are the second most popular data integration tools types after ETL tools”

*Gartner's 6/6/08 report “Survey on Data Integration Practices Shows Move Toward Strategic Initiatives”*



# Informatica Data Replication

## Product Overview



# What is Data Replication

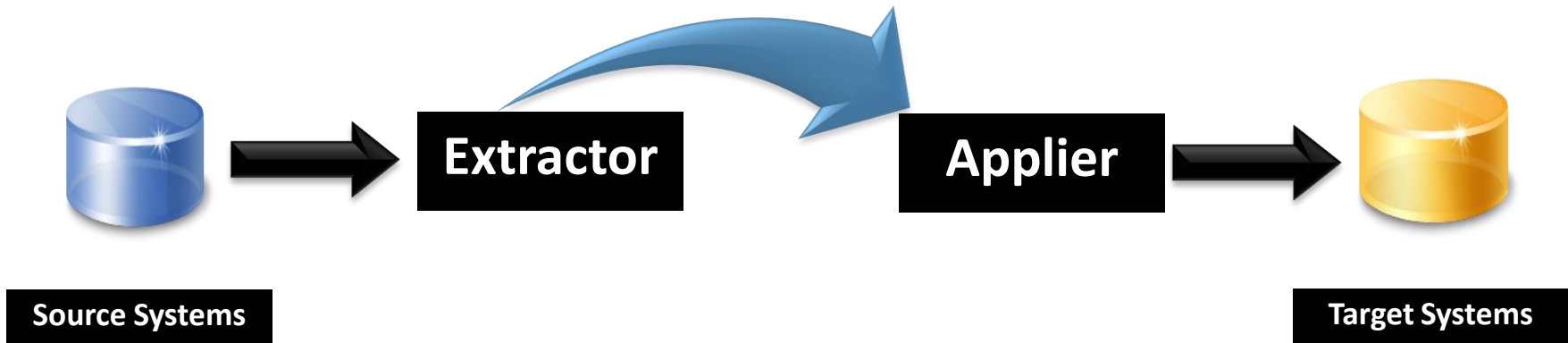
A set of technologies to

- manage the capture, routing and delivery of transactional data in **real-time**
- **automate** the **cloning** of application **data** – thousands of mixed tables at once



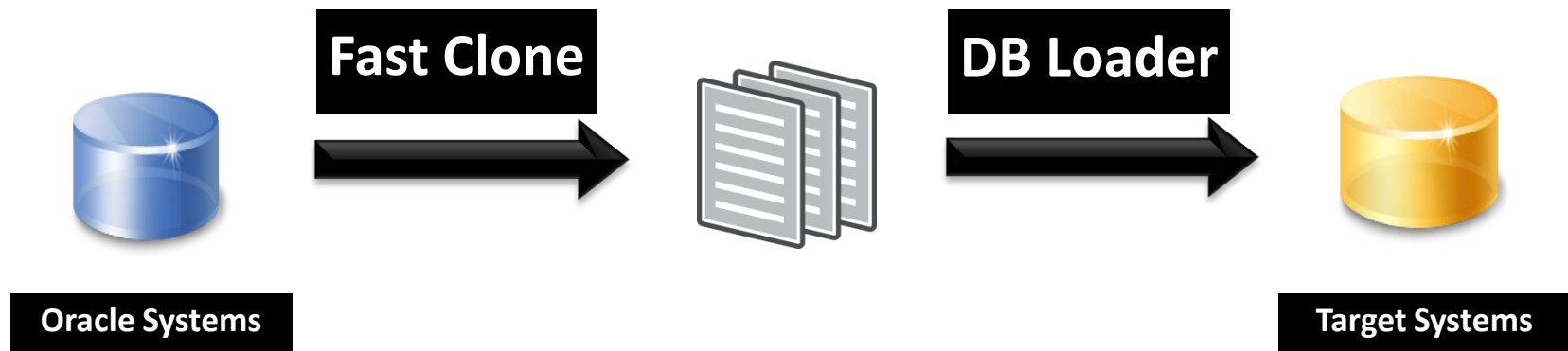
# What is Informatica Data Replication

Captures, routes and delivers transactional data across different systems (consisting of thousands of tables and hundreds of schemas) in real-time, while maintaining order and consistency, and with minimal impact to the source systems

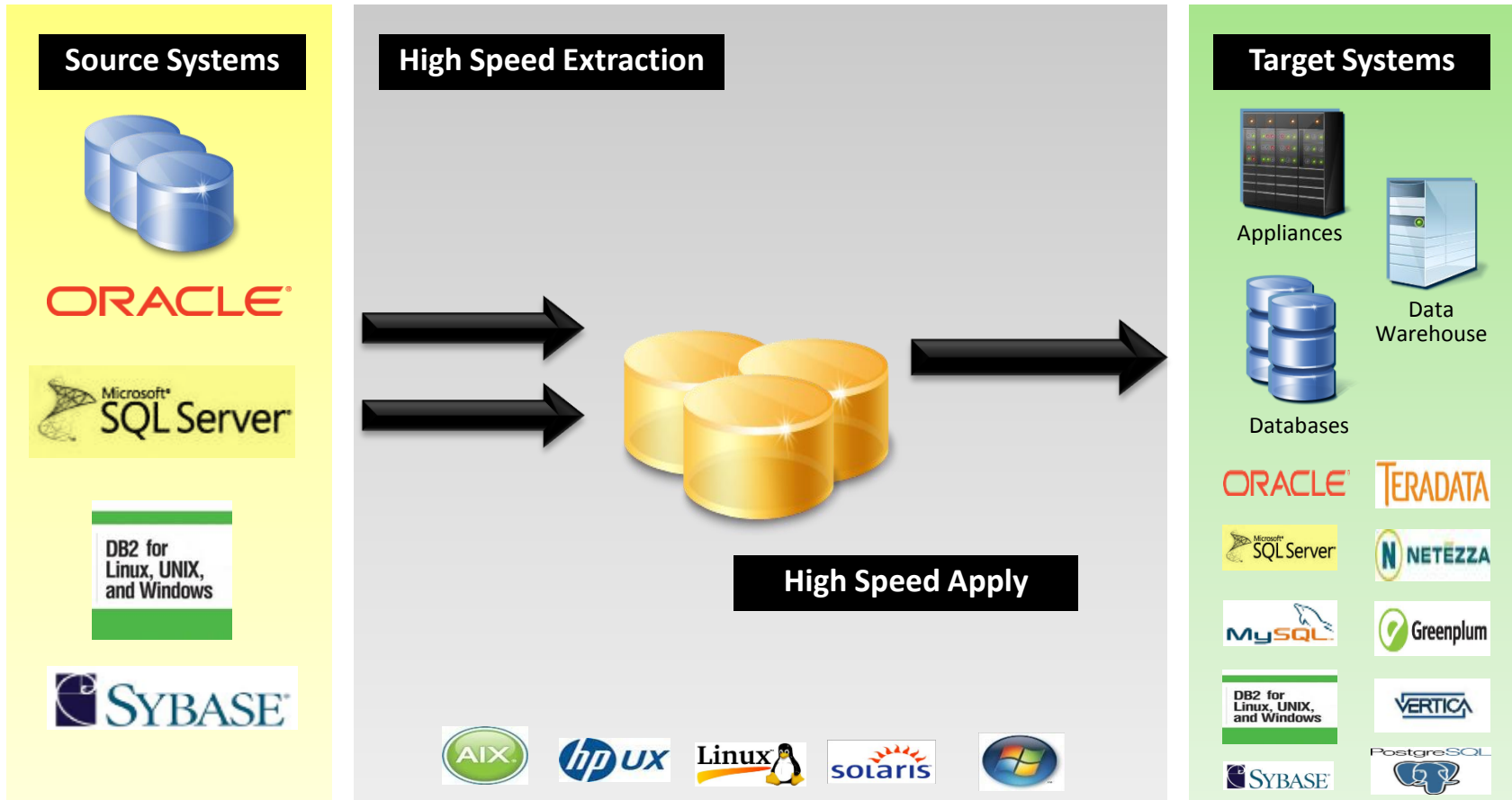


# What is Informatica Fast Clone

High speed extraction from Oracle and loading into target databases and appliances



# Informatica Data Replication Solution



**Data Replication:** Replicate Transaction-oriented incremental changes from source to target

**Fast Clone:** copy data directly from Oracle to a supported target, in a batch mode process

# Informatica's Data Replication

## The benefits of Data Replication in our environment

<b>Supports Active Data Warehouses</b>	✓ Ensures Fresher, Up-to-date information from business critical online systems
	✓ Leverages investment in potential existing 'ETL' data warehouse approach
<b>Enables real-time Reporting with Zero Impact to the Production Systems</b>	✓ Avoid Performance impact on your operational Database
	✓ Get Real-Time Data Capture
	✓ Up-to-date information for critical business systems
<b>Provides Uninterrupted Migration Path</b>	✓ Enables movement from an older OS or database platform version to a new one with no interruption to operations
<b>Live Auditing</b>	✓ Reduce compliance risks with full audit trail

# Informatica Data Replication Solution

## Competitive Differentiators

### Fast Time to Value

Quick configuration and rapid deployment via easy, comprehensive GUI  
Single day versus weeks to deploy

### Heterogeneous Platform Support

Architected for transactional replication across many source and target  
database and OS platforms

### Optimized for Real Time Data Warehousing

Uses innovative techniques to optimize transaction loading into warehouses

### Live Audit Capability

Audit apply allows full transaction auditing without impacting production system

# Informatica Data Replication

Solutions | Use Cases

# Key Solution Areas

## Real Time Data Warehousing for BI and Analytics

Building near real-time operational reporting system with minimal impact on production  
Optimized Apply processes for Data Warehouse (Teradata, Netezza, Greenplum, Vertica)

## Operational Reporting

Build near real-time operational reporting system with minimal impact on production

## Query Offloading

Distribute Resource Utilization creating a cost effective standby system for query / reads

## Continuous Availability

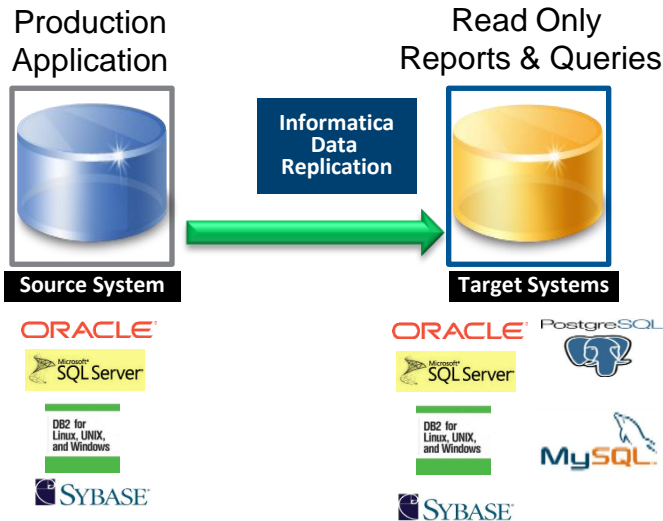
Minimal Downtime Migrations & Upgrades  
Migrate Database, Hardware, O/S, or Application with minimal downtime



# Query Offloading

## Infrastructure Cost Reduction

Reduce Cost  
Increase Productivity



### Solution

- Create a real-time replica of existing expensive operational system to offload reporting and query activity.
- Sync replica in real-time and route all read-only traffic

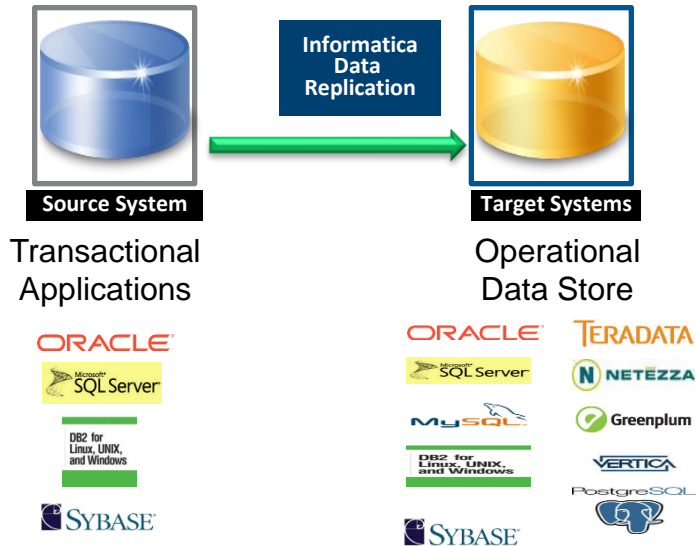
### Benefits

- Reduce/Eliminate query load on transactional system
- Directly reduce cost associated with additional licensing fees or the need for specific Database enterprise tools

*If you want to move query processing to standard, lower cost DB & Hardware infrastructure with no disruption to critical transaction processing systems & satisfy business end-users as data volumes and users grow you should consider Informatica Data Replication (IDR)*

# Real Time Data Integration

## Operational Reporting



### Solution

- Real-time analytics from transactional systems
- No additional load on source system – offload reporting to a database instance that is synchronized in real time

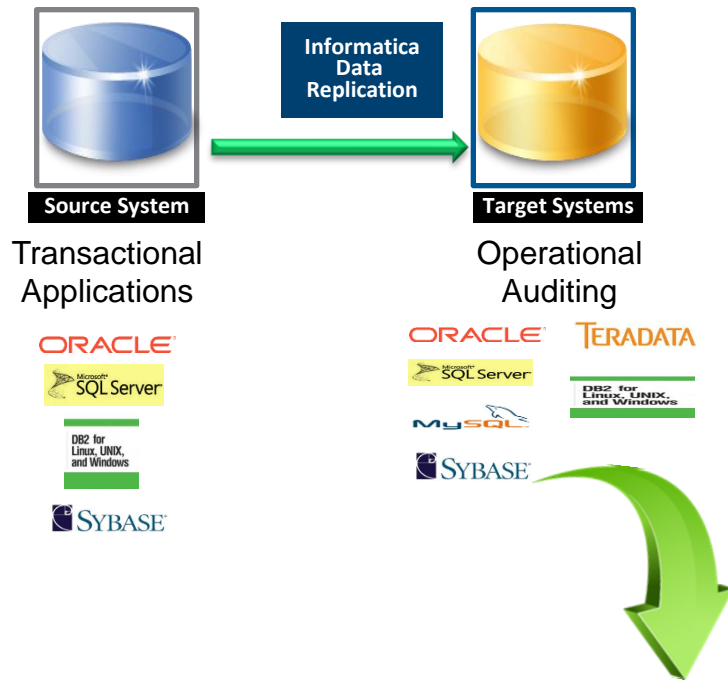
### Benefit

- Enables tactical decision making and business transparency
- Provide 360 degree views of your customers, your business, for trend analysis
- Offloads impact from source systems
- Building a single view of the truth and for true real-time analytical analysis
- Built in optimization for Data Warehouse, Data Warehouse Appliance delivery
- Compliments ETL solutions, provides integration of existing Informatica

*If you need real-time analytics for operational Business Intelligence you need to consider Informatica Data Replication (IDR)*

# Optimized Integration Alternatives

## Real-time Transaction Auditing



### Solution

- Real-time auditing of transactional systems to meet compliancy regulations
- Build, Track, and Maintain a complete or selective audit record of your OLTP system.

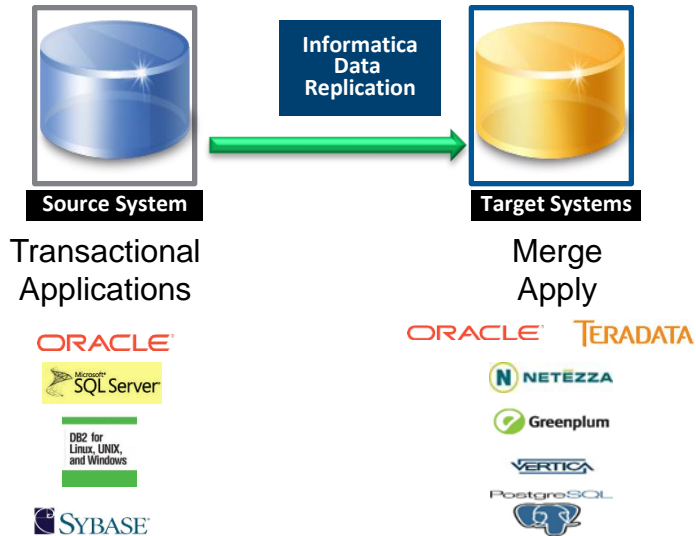
### Benefit

- Enables real-time auditing capabilities across all or selective tables in a schema / database
- Maintain activity audit records for compliancy
- Identify transactional patterns based on user activity
- Develop a fast alert tracking system for fraud detection
- Flexible mapping options allows support for multiple apply methods in a single configuration
- Built in table creation function automates the creation of Audit \_LOG tables

OP_CODE	OP_TIME	OP_CMT_SCN	OP_CMT_TIME	OP_XID	ID_NEW	ID_OLD	NAME_NEW	NAME_OLD	ADDRESS_NEW	ADDRESS_OLD
I	2008/05/19 13:54:16	160754728	2008/05/19 13:54:20	1970359196746518	2		Kyle Brown		875 Paulson St., Bothell 98012, WA	
U	2008/05/19 13:54:28	160754733	2008/05/19 13:54:31	1688922874745326		2			123 Preston St., Bothell 98013, WA	875 Paulson St., Bothell 98012, WA
D	2008/05/19 13:54:37	160754737	2008/05/19 13:54:38	2251997382217908				Kyle Brown		123 Preston St., Bothell 98013, WA

# Optimized Integration Alternatives

## Optimized Loading of Data Warehouse Appliance



### Solution

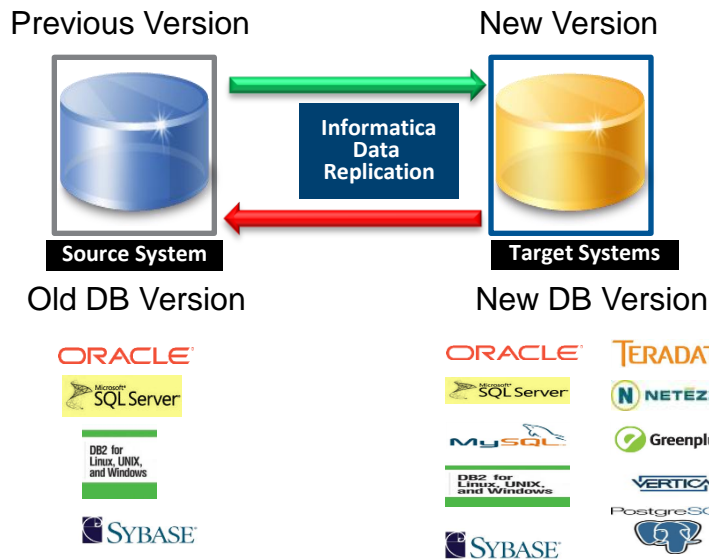
- Loading Data Warehouse Appliances for testing and validation
- Continuous real-time feeding of Data Warehouse Appliances

### Benefit

- Optimized Merge Apply process for efficiently loading Data Warehouse Appliances
- Rapid deployment of Data Warehouse Appliances from Test to Production
- Automatic Mapping from heterogeneous sources
- Parallel Apply processing to address the most demanding high volume environments
- Specifically designed and optimized for Data Warehouse Appliances
- Built in table creation function automates the creation of Audit \_LOG tables and Merge Tables.

# Continuous Availability

## Reduced Downtime for Migrations and Upgrades



### Solution

- Migrate from any heterogeneous database to IDR supported targets with minimal downtime
- Migrate hardware or OS platforms with no downtime
- Upgrade Applications with no downtime
- Replicate from new version back to old to extend the fail-back window

### Benefit

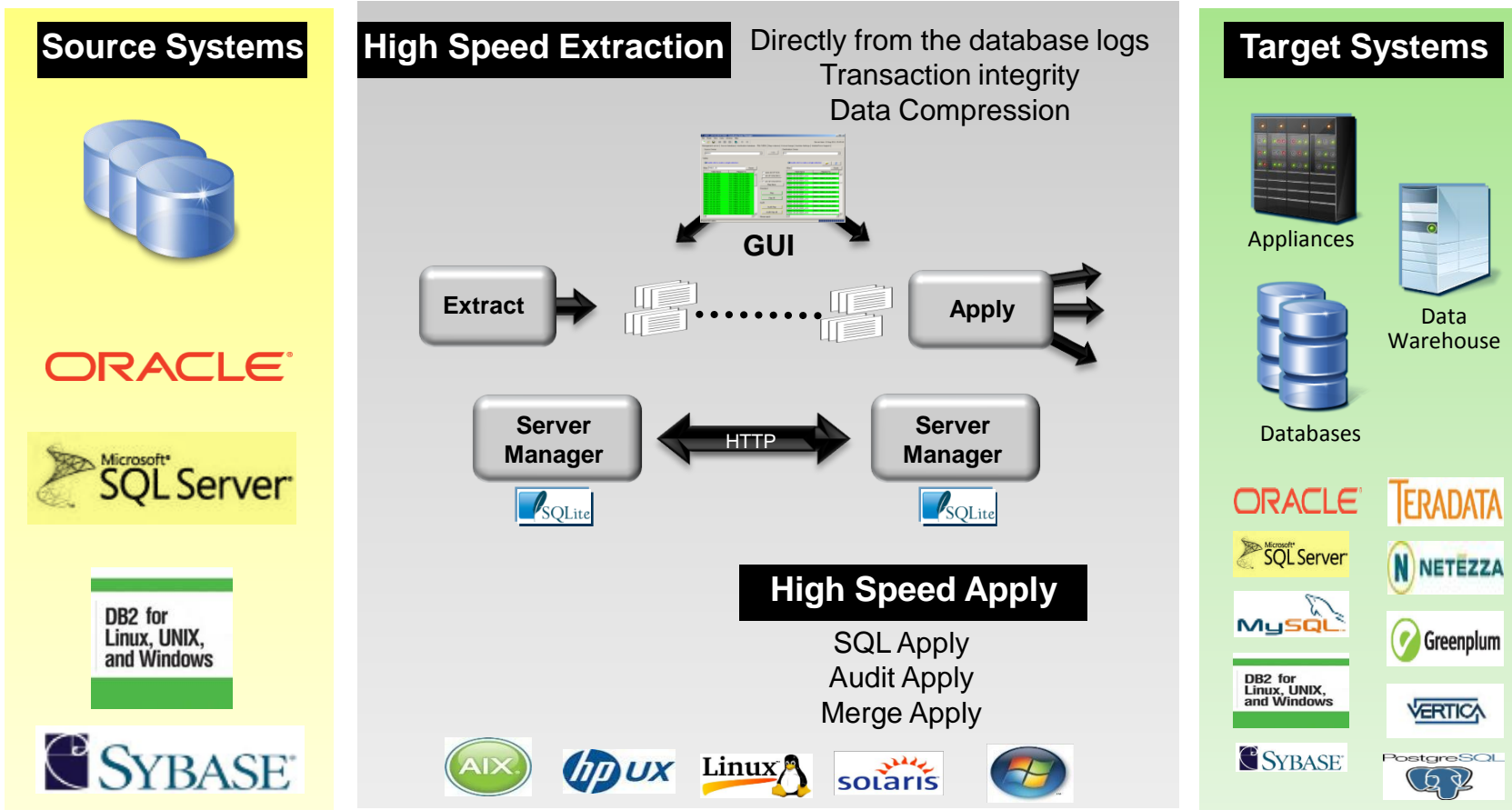
- Standardize your data center and modernize platforms
- Zero down-time for mission critical applications
- Migrate from one system to the next, while keeping production system live.
- Test new system with active data from product
- Lower risk by enabling failback

*If you need to migrate your database or upgrade critical applications but can't afford to take any down time and impact critical business processes Informatica Data Replication can help!*

# Informatica Data Replication

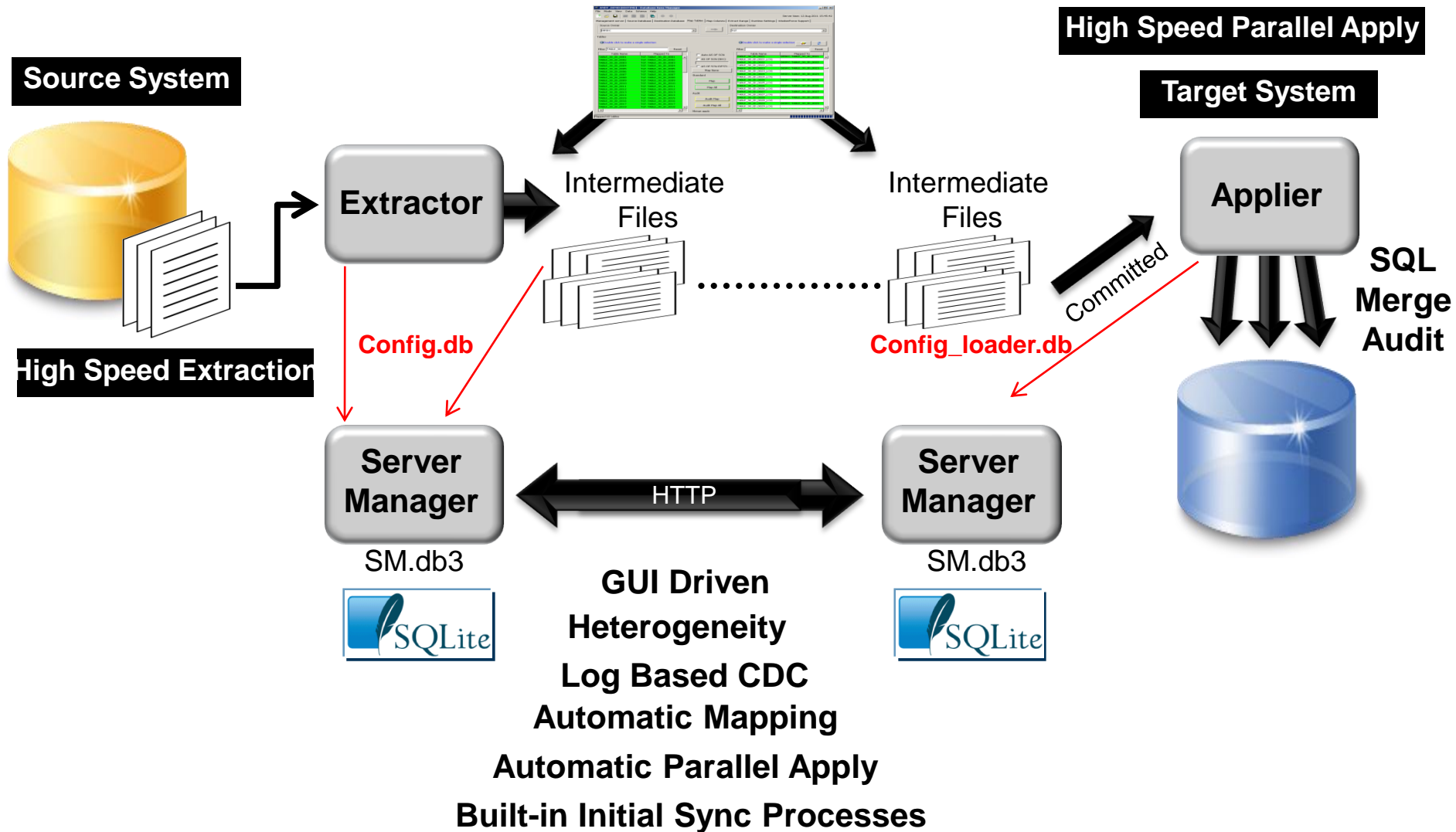
## Architecture

# Informatica Data Replication Architecture



Synchronize sources to targets  
Switch to Replication

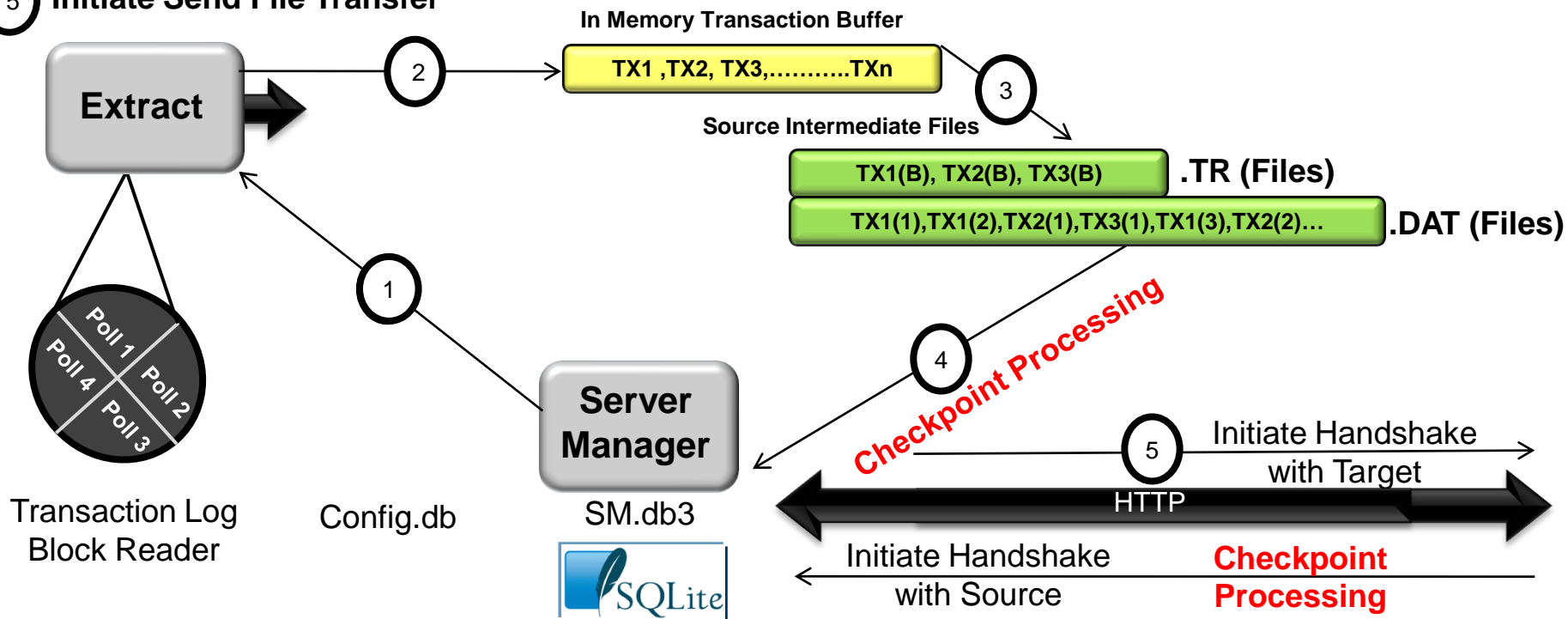
# Informatica Data Replication Solution





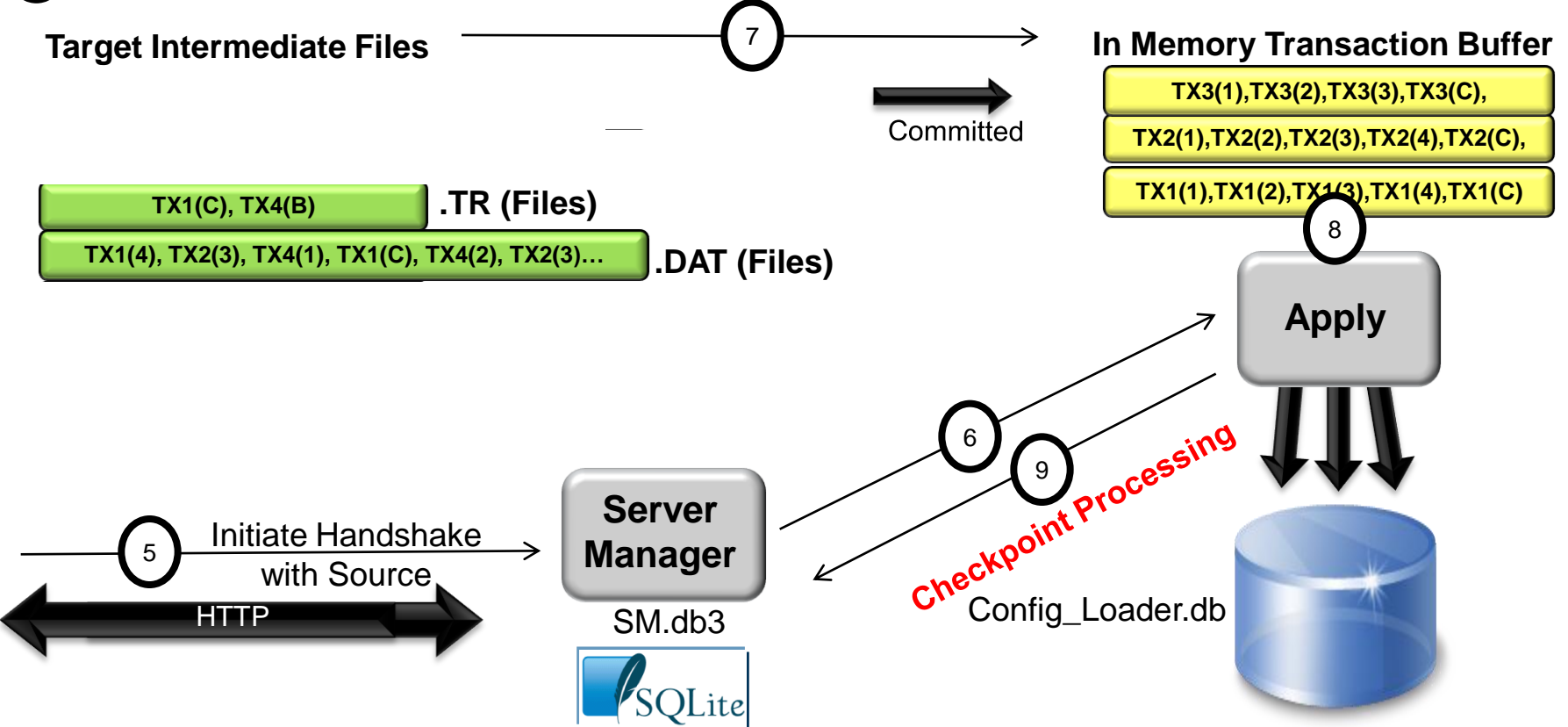
# Informatica Data Replication : Extract

- 1 Request to begin Extract is sent from the Main Server Manager
- 2 Extract begins reading from the transaction logs, building an In-Memory Queue
- 3 During a polling cycle, the buffer is flushed to the intermediate files (.TR/.DAT)
- 4 Process is checkpointed
- 5 Initiate Send File Transfer



# Informatica Data Replication : Apply

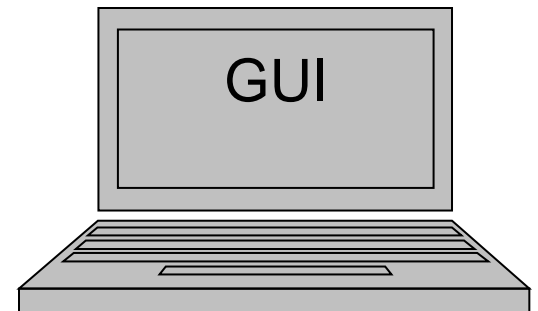
- 6 Request to begin Apply Process
- 7 Apply begins buffering only committed transactions from Intermediate Files
- 8 Apply Process delivers committed transactions across multiple threads, then commits
- 9 Process is checkpointed, and unnecessary Intermediate Files are automatically pruned



# Informatica Data Replication

## *Component: IDR Console*

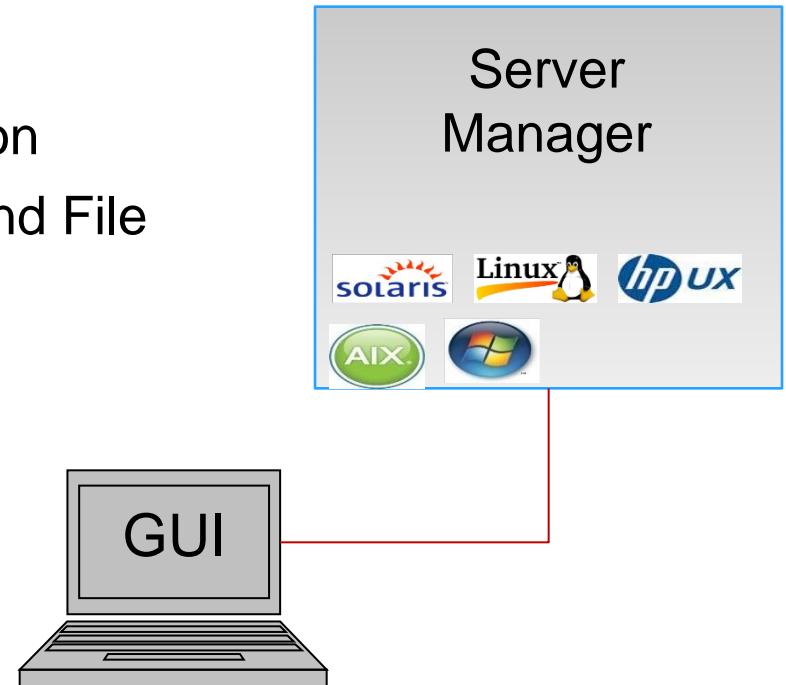
- Windows / X-Windows GUI interface
  - Define local/remote servers
  - Configure replication criteria
    - Define source and destination databases
    - Configure table and column mappings
    - Choose Apply method
  - Schema DDL Conversion
  - Define Schedules
  - View Reports
  - Set Alerts



# Informatica Data Replication

## Component: Server Manager

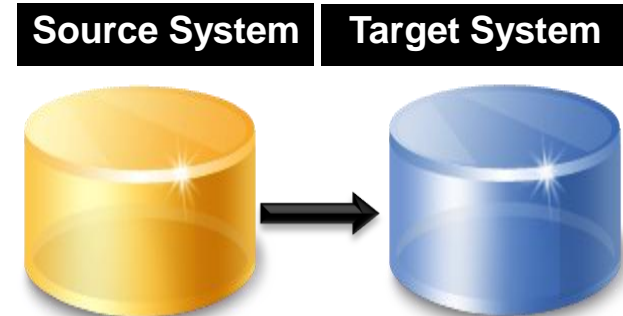
- Main Server
  - Task management
    - Start the InitialSync task
    - Start and monitor Extract/Apply tasks
    - Handles the Send File task
    - Maintains task status information
  - Deletes Intermediate Files after Send File
  - Schedule and monitoring jobs
  - Maintain logs
- Subserver Server
  - Task management
    - Start InitSync, Extract, Apply
    - Handles the Send File task



# Informatica Data Replication

## Component: InitialSync

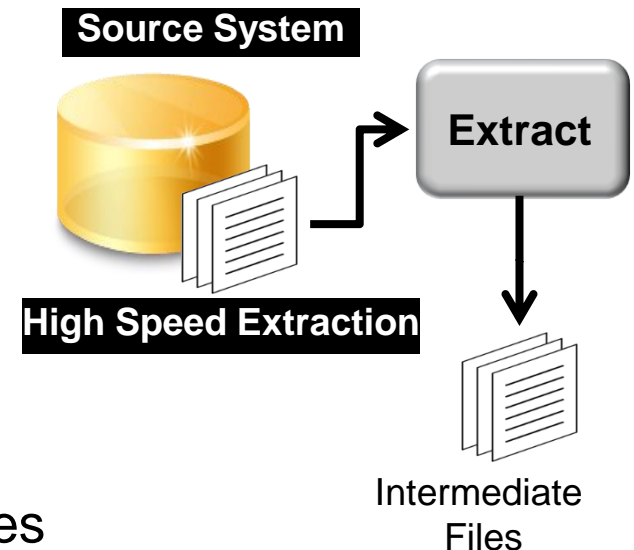
- Synchronizes existing data in source tables with target tables
  - Parallelization ensure very high performance
  - Works across database platforms
  - Method used:
    - Oracle: DBlinks or Direct Load API
    - SQL Server: BCP
    - Greenplum: LibPQ Copy Command
    - Teradata: Teradata Parallel Transporter
    - Netezza: Remote External Tables



# Informatica Data Replication

## *Component: Extract*

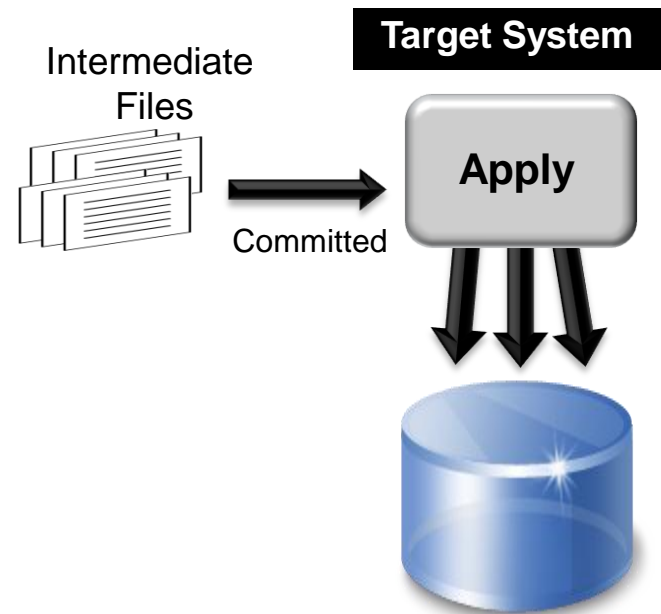
- Parses the transactional log files and extracts the relevant transactional data
- Changes are stored in compressed on-disk intermediate files
- High performance and low resource usage
- Transactional changes are extracted from:
  - Oracle online redo or archive logs
  - DB2 logs
  - SQL Server transactional journal
- Can parse logs produced on other architectures



# Informatica Data Replication

## *Component: Apply*

- Applies the committed changes from the intermediate files into the target database
- High performance achieved through parallelization and bulk loading
- Three Apply Methods:
  - SQL Apply (pure replication)
  - Audit Table Apply
  - Merge Apply





**INFORMATICA**<sup>®</sup>  
The Data Integration Company<sup>™</sup>