IBM Software Group

# DB2 Test Database Generator for z/OS Version 2.1

2 Février 2005

Catherine Chochoy (source: eBU presentation by Ed Lynch, 2004)

DB2 Information Management Software

e business on demand software



## Agenda

- Test Database Generation Objective
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- Hints & Tips

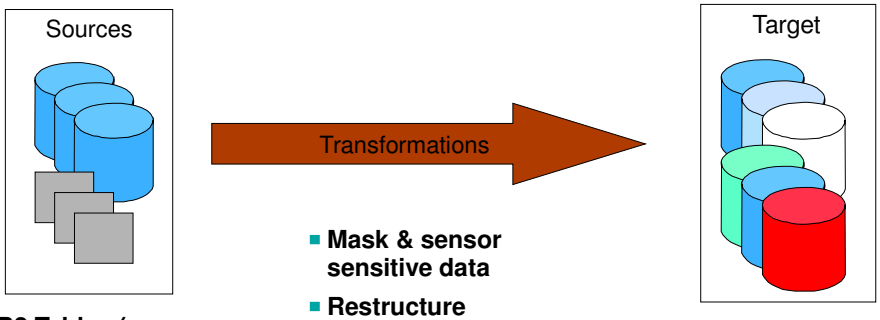
TDBG = Test Database Generator

# Agenda

- **Test Data Generation Objectives**
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- Hints & Tips



# Test Data Generation: Customer Requirements




- **DB2 Tables (any platform)**
  - **Recognize relationships**
- **Text Files (e.g. fixed, delimited, XML)**
- **Non-existent sources**

- **Mask & sensor sensitive data**
- **Restructure objects**
- **Filter**
- **Slice of related data**
- **Generate data from scratch**

- **SQL**
- **Direct to DB2 Tables (any platform)**
- **Text Files (e.g. fixed, delimited, XML)**
- **Load Format**
- **New or Existing**

# Agenda

- Test Data Generation Objectives
- **IBM DB2 Test Database Generator for z/OS Version 2.1 Today**
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- Hints & Tips



# IBM DB2 Test Database Generator for z/OS Version 2.1 (Today)

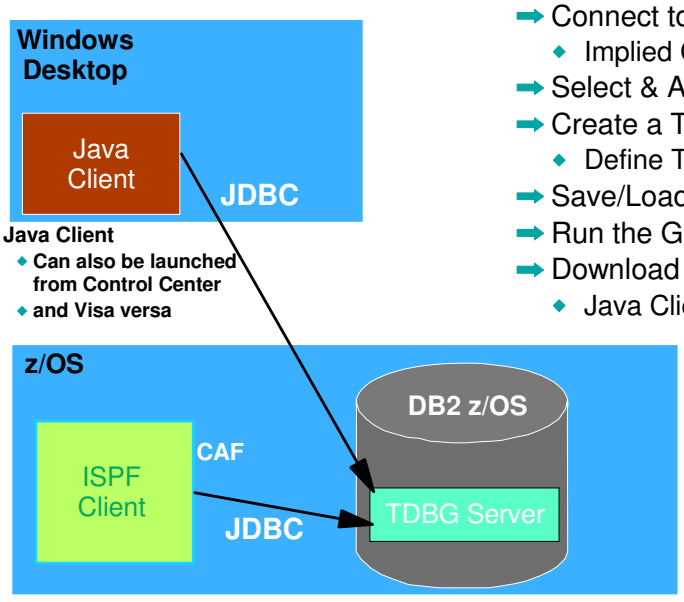
- Sources:
  - DB2 for z/OS Table(s)
    - All Rows
    - Every nth row
    - Random
    - Predicate Filter
  - Delimited/Fixed Text File
    - Local File
    - FTP site
    - HTTP site
- Relational Relationships (Grouper)
  - DB2 Defined RI
  - Application RI
- Choice of Tool Interface
  - Java Client (Windows)
  - ISPF
- Transformations
  - Columns (new, order, name)
  - Data Attributes - type, lengths,nulls
  - Filter
  - Rules:
    - Source Column
    - Static
    - Lookup
    - Mask
    - Expression
    - Random
    - Pattern
- Target Output Formats:
  - Limit Number of Rows
  - DB2 Tables (z/OS & Multiplatform)
  - SQL
  - Delimited/Fixed Text File
  - XML
  - DB2 for z/OS Load Format
- Transferable Outputs

# Agenda

- Test Data Generation Objectives
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- **Terminology and Architecture**
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- Hints & Tips



# TDBG Client



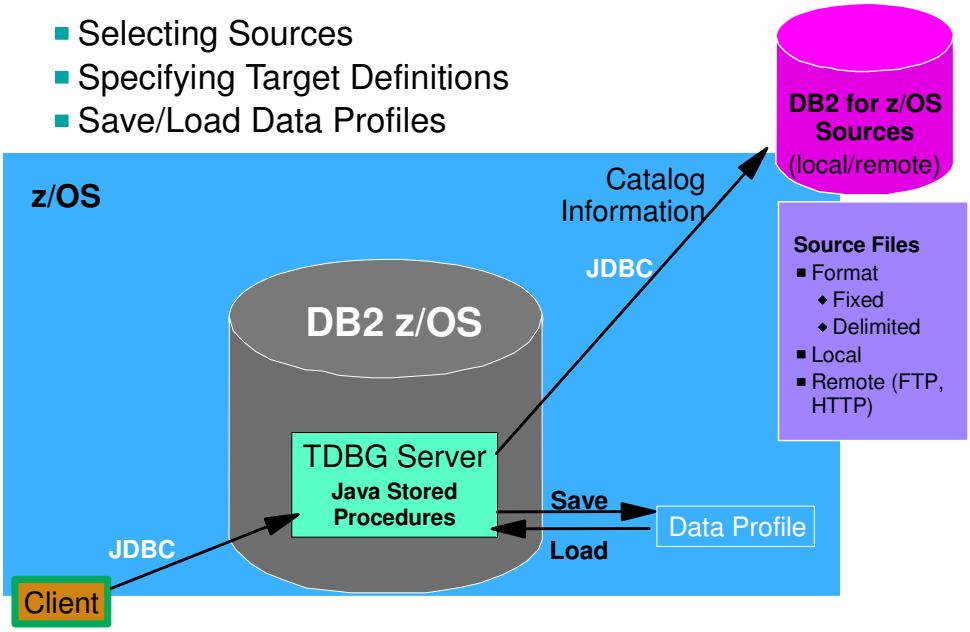

**Java or ISPF Client - Same Capabilities**

- Connect to TDBG Server
  - Implied Connect for ISPF client
- Select & Add Source(s)
- Create a Target
  - Define Transformations
- Save/Load Data "Profiles"
- Run the Generator
- Download Generated Data
  - Java Client only

JDBC Drivers Supported
- Type 2 or
- Type 4

# TDBG Server and Sources

- Selecting Sources
- Specifying Target Definitions
- Save/Load Data Profiles



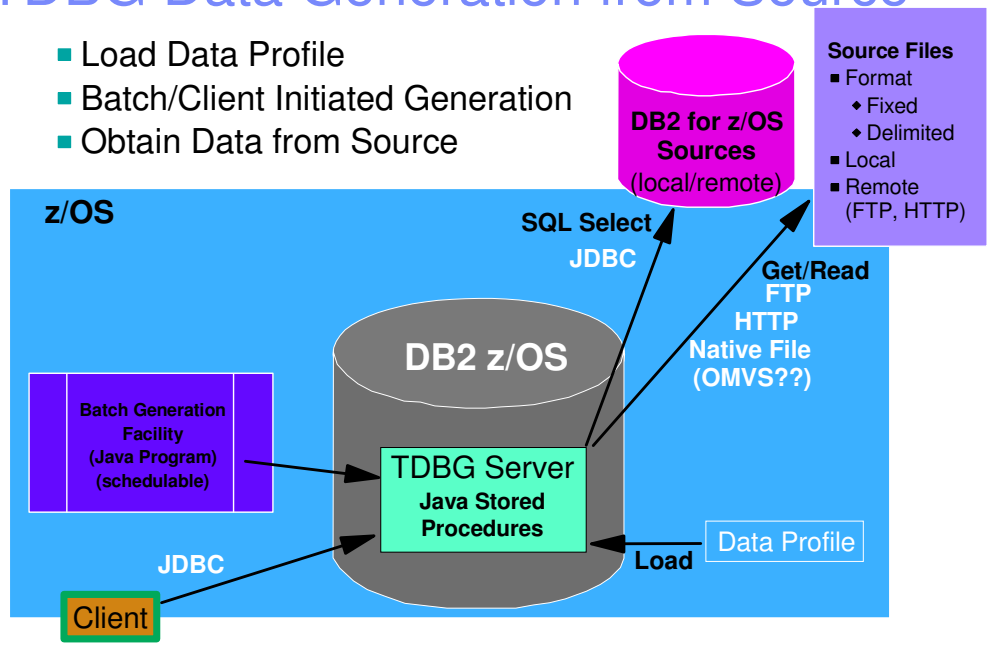



# Data Profiles

- Data Profiles define the generation process
  - Source data objects used to seed generation
  - Definition and Relational Rules of the Targets
  - Describes how data will be copied, filtered, masked, and/or transformed
- Data Profiles
  - Created by the TDBG Server and
  - Stored in the HFS where the Server is running
- Data Profiles are written in Test Database Generator Markup Language (a.k.a. GRIML)
  - GRI: Internal product code
  - ML: Markup Language
  - GRIML is an XML-based markup language
- Can manually create using an XML Editor

# TDBG Data Generation from Source

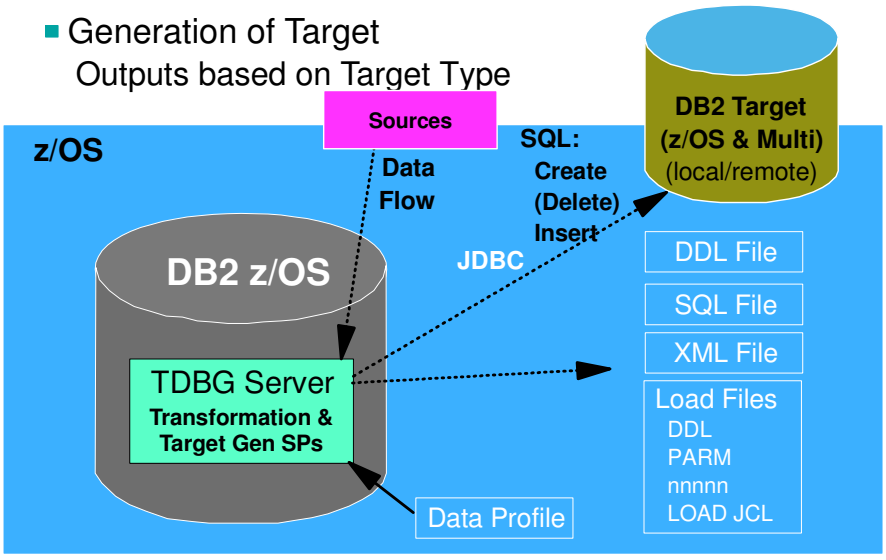
- Load Data Profile
- Batch/Client Initiated Generation
- Obtain Data from Source





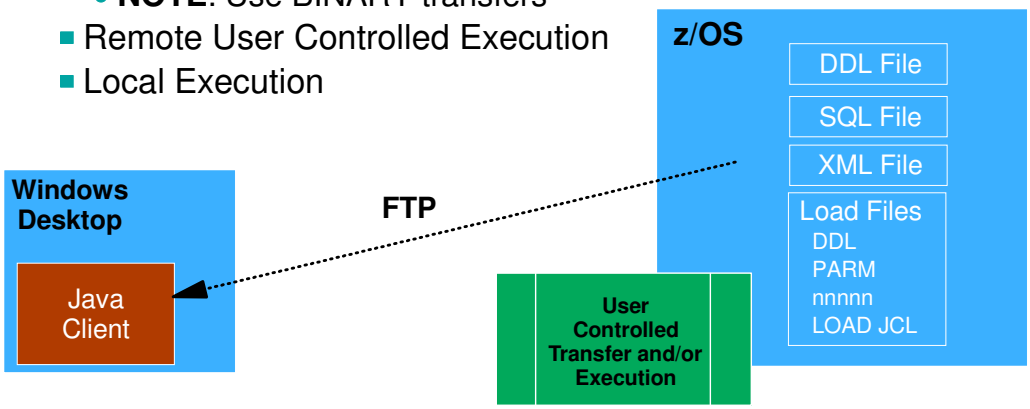

# TDBG Generation of Target

- Transformation of Data
- Generation of Target
  Outputs based on Target Type



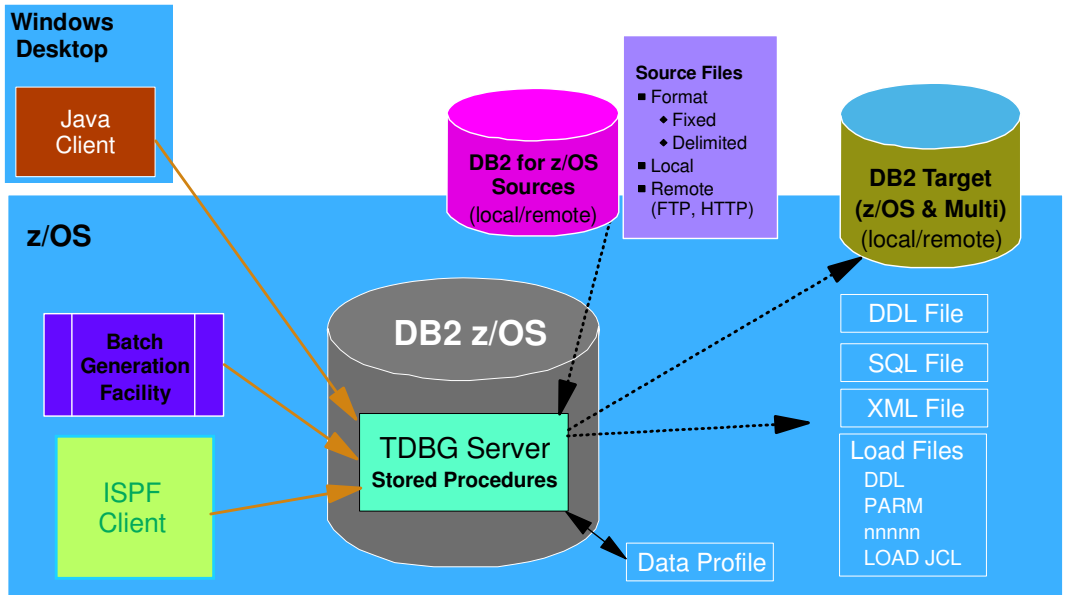
# Using TDBG Outputs

- All output is UNICODE UTF-8
- Download Output
  - Using Java Client to Local or Network Disk Drive
  - User controlled file transfer (e.g. FTP)
  - **NOTE**: Use BINARY transfers
- Remote User Controlled Execution
- Local Execution






# TDBG Overall Architecture



# Agenda

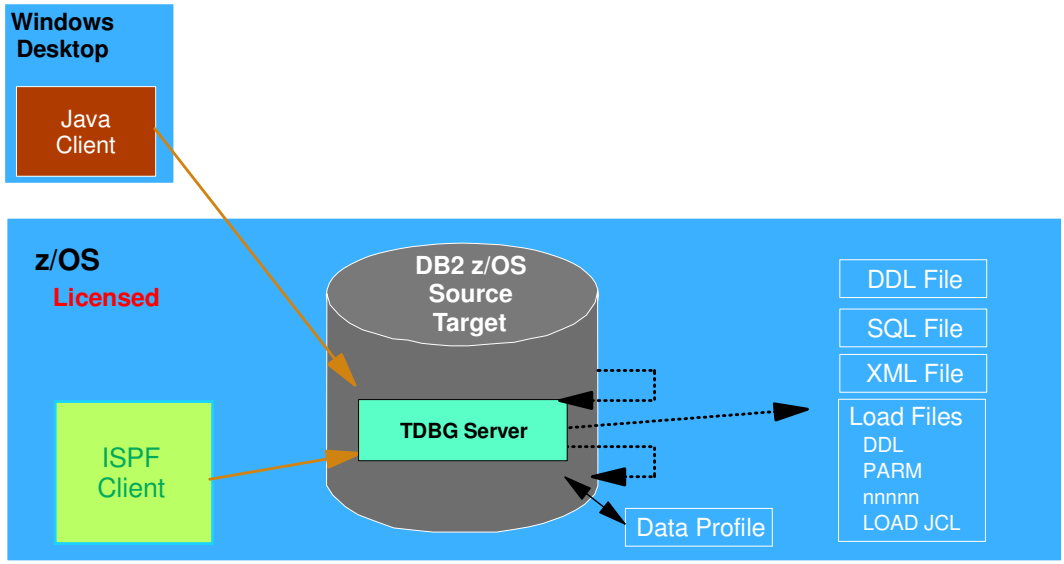
- Test Data Generation Objectives
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- **Potential Configurations**
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- Hints & Tips



# Licensing

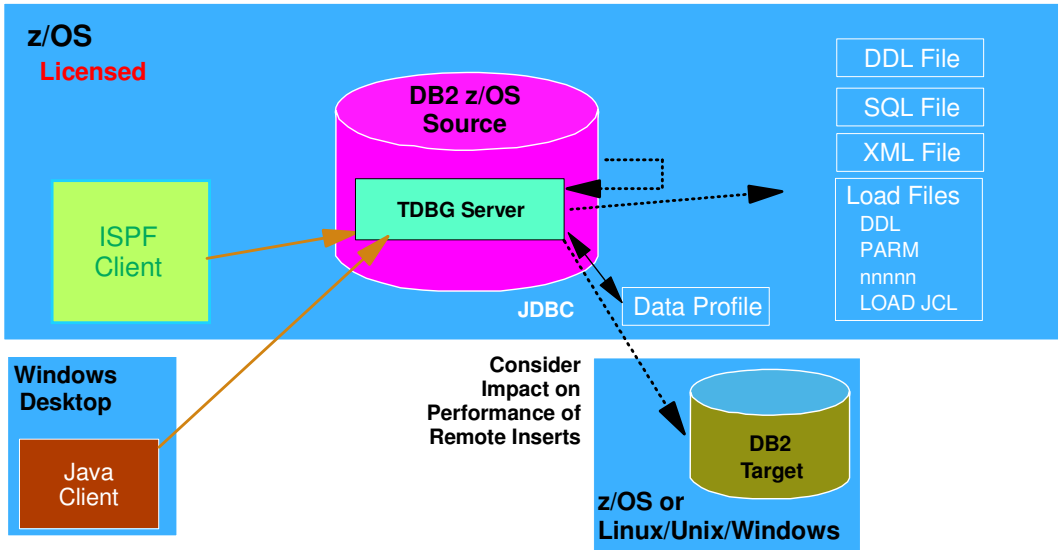
- Charges for Licensing of DB2 Test Database Generator for z/OS is based upon:
  - Processor where the TDBG Server is Located, and
  - Processor(s) where the Source(s) are Located

# Combined Source/Target/Server

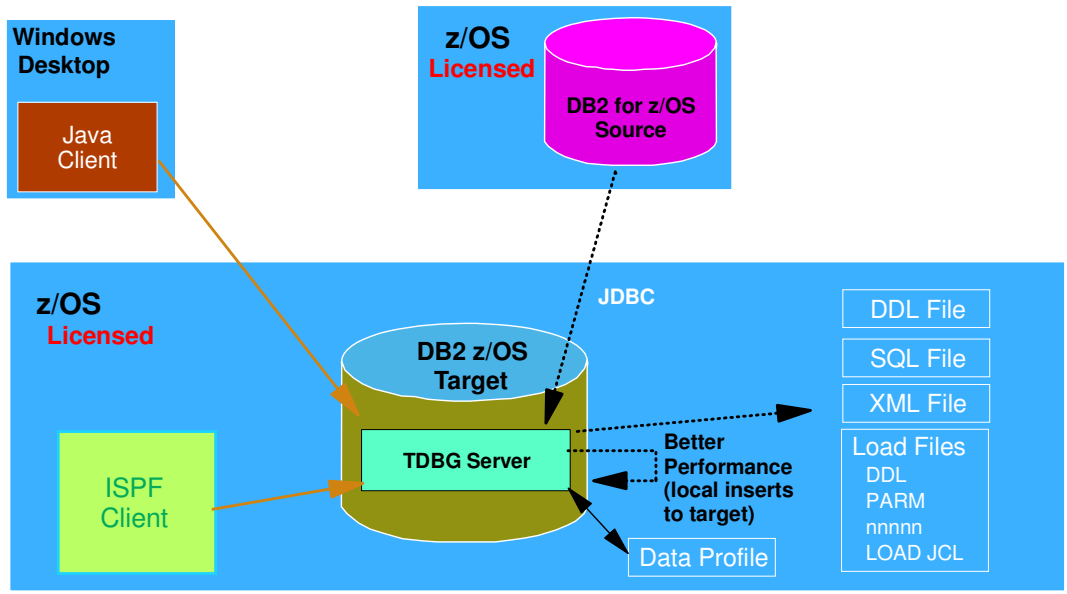





# Combined Source & Server with Remote Target (any platform)

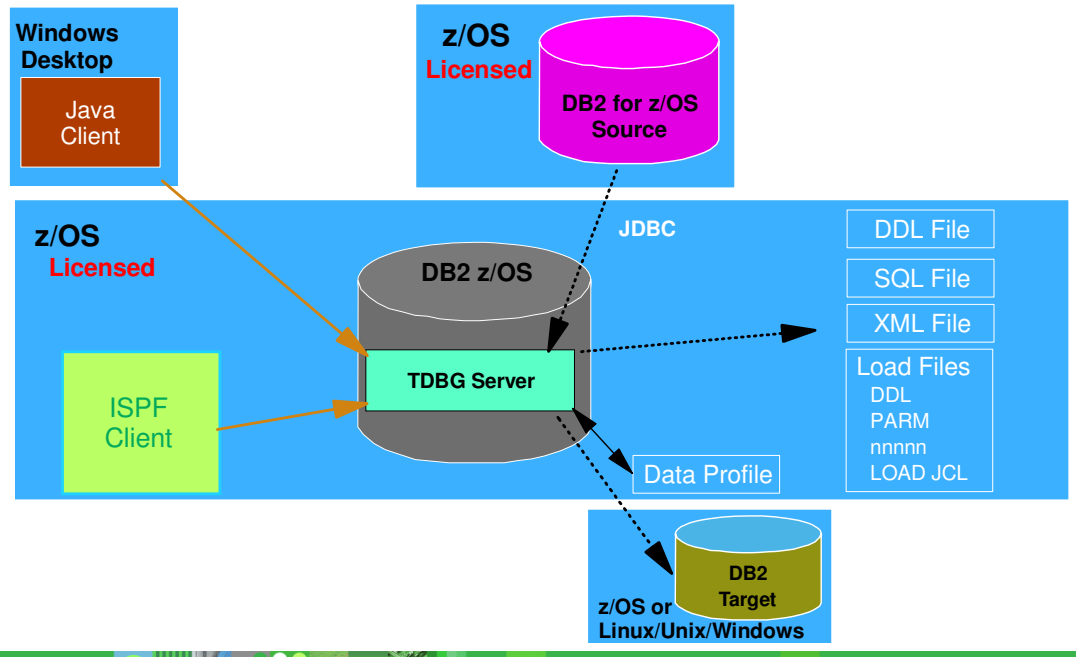


# Combined Target & Server with Remote Source






# Separate Source, Target, & Server



# Agenda

- Test Data Generation Objectives
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- **Transformations**
  - **Source & Target Specifications**
  - Transformation Rules
- Generation & Results
- Hints & Tips



# Data Transformation Capabilities

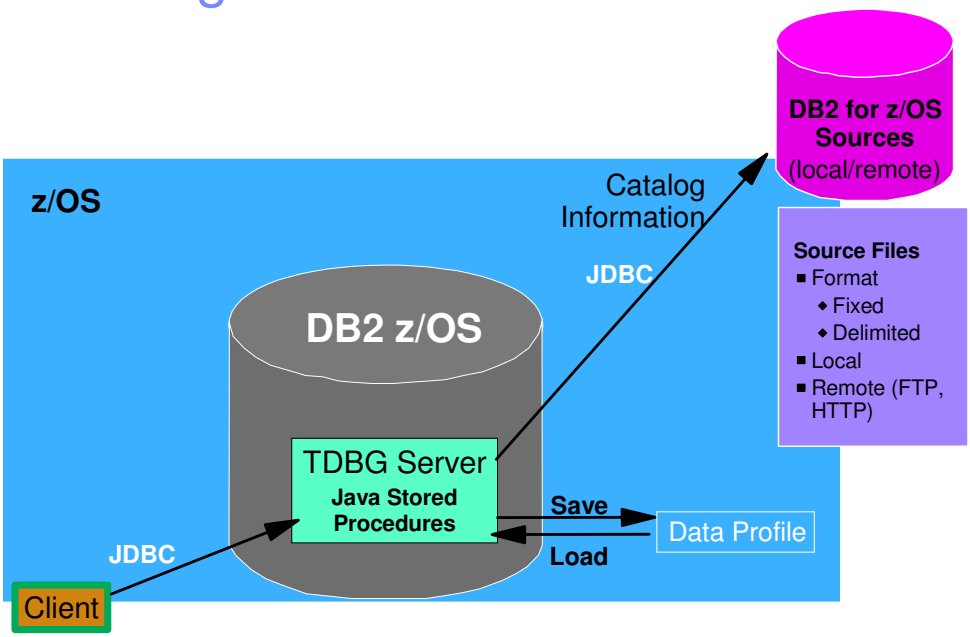
- Table
  - Change Table name
  - Change Table Creator
- Columns
  - Add (create) / Exclude
  - Name Change
  - Order of Columns
  - Specify / Change Data Attributes
    - Data Type
    - Length(s)
    - Nullability
  - SQL Transformations
  - Specify Transformation Rule(s):
    - Source Column
    - Static
    - Lookup
    - Mask
    - Expression
    - Random
    - Pattern
- Row Controls
  - Filtering via SQL Predicate
  - Row Range
    - All
    - From x to y
  - Selection within the Range
    - All
    - Every nth Row
    - Random
  - Generate data from scratch
  - Control Number of Rows Generated
- Join / Union Source Tables
- Single Source to Multiple Unique Targets
- Related Tables (Requires Grouper)
  - Identify and Include
- Target DDL Generation

# Source & Target Specifications

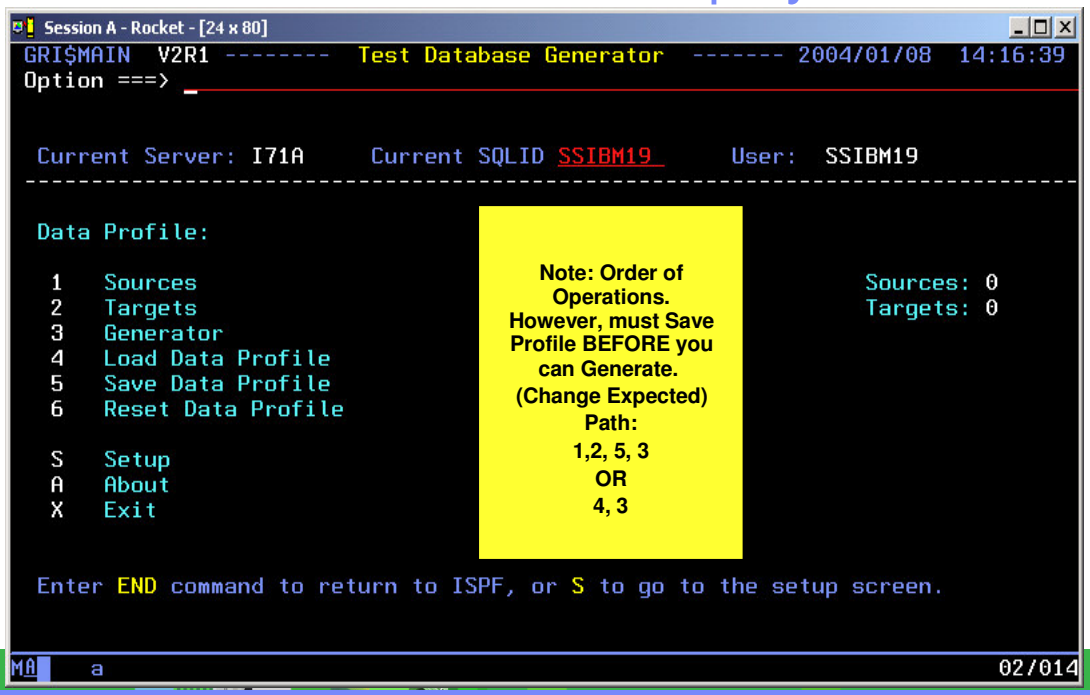
- Data Transformations are Specified
  - Using the TDBG Client Interface
    - With Either the Java or ISPF Client
  - While Specifying your Source Object(s) and
  - While Creating your Target Object(s)
- Start Your Client
  - Connect to the Server
    - Explicitly for Java
    - Implicit with ISPF
      - CLIST Invocation Specifies DB2 Subsystem



# Connecting to a TDBG Server



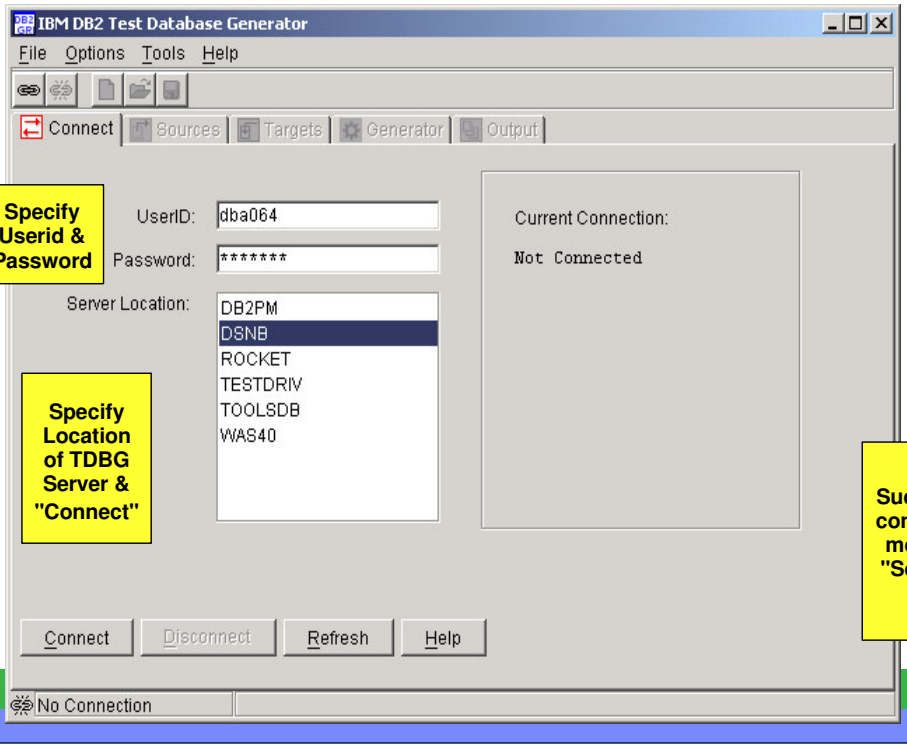
# Initial TDBG ISPF Client Display




**Note: Order of Operations. However, must Save Profile BEFORE you can Generate. (Change Expected) Path: 1,2, 5, 3 OR 4, 3**



# Connect to TDBG Server (Java Client)

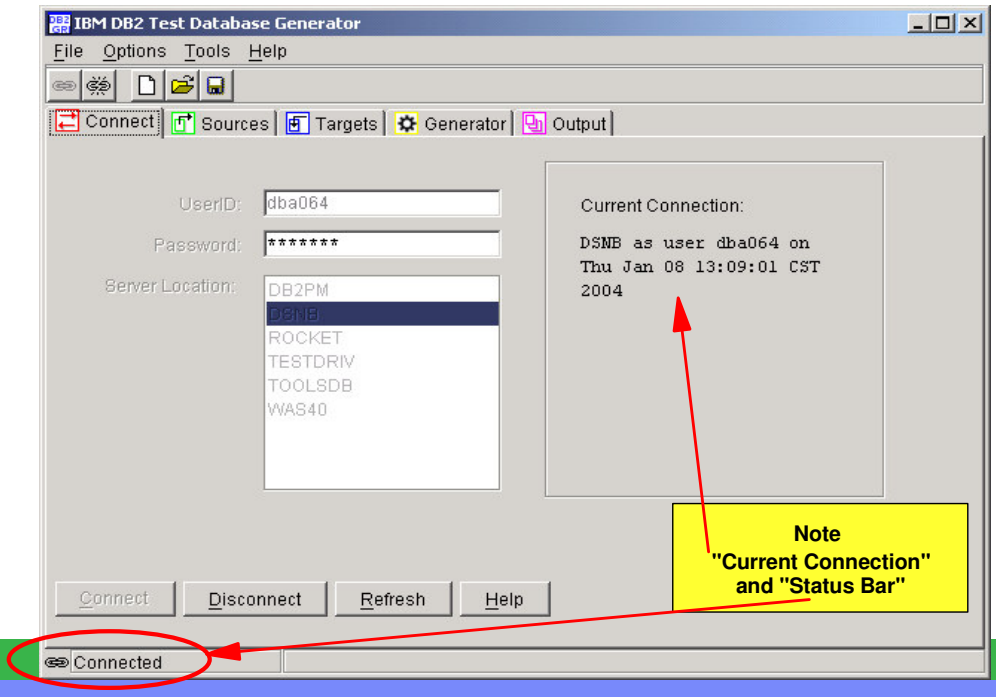



**Specify Userid & Password**

**Specify Location of TDBG Server & "Connect"**

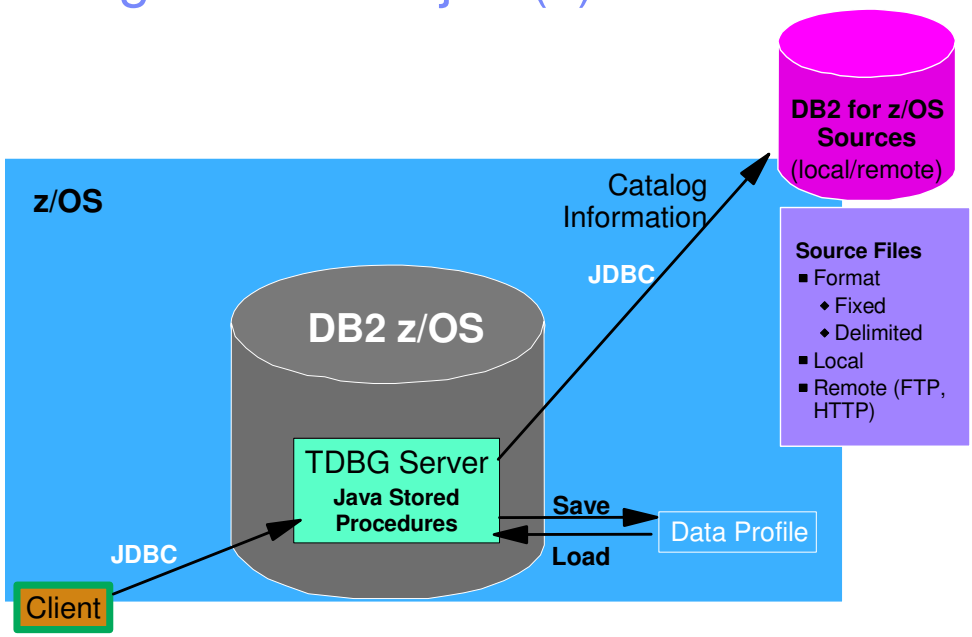
**Successful connection moves to "Sources" tab**

# Successful TDBG Connect Display

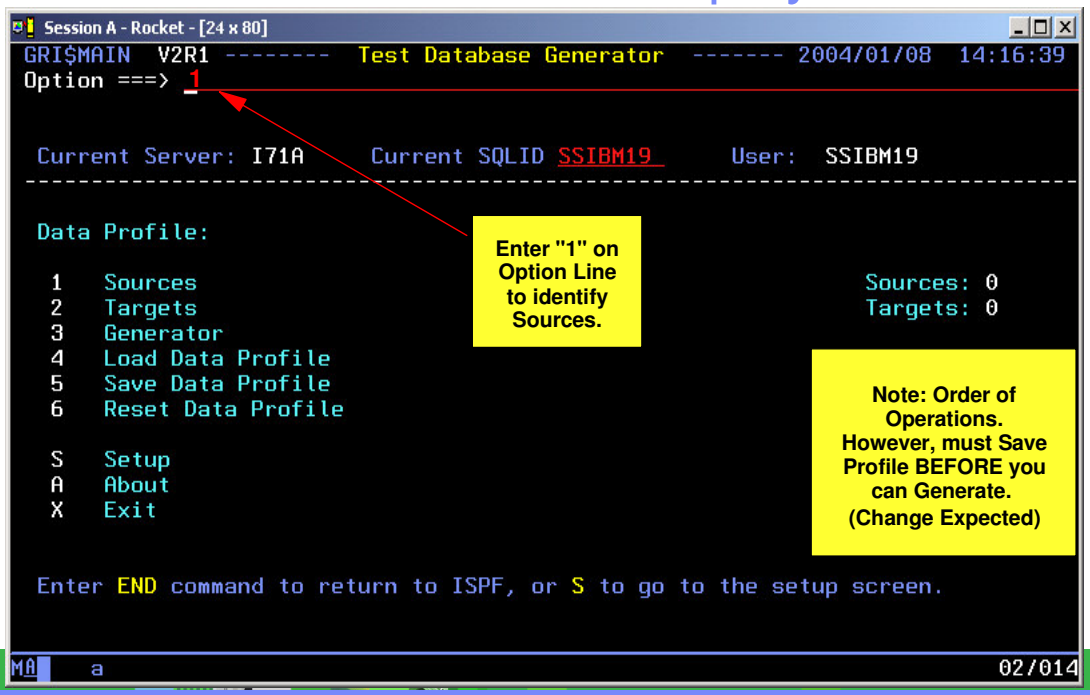





# Defining Source Object(s)

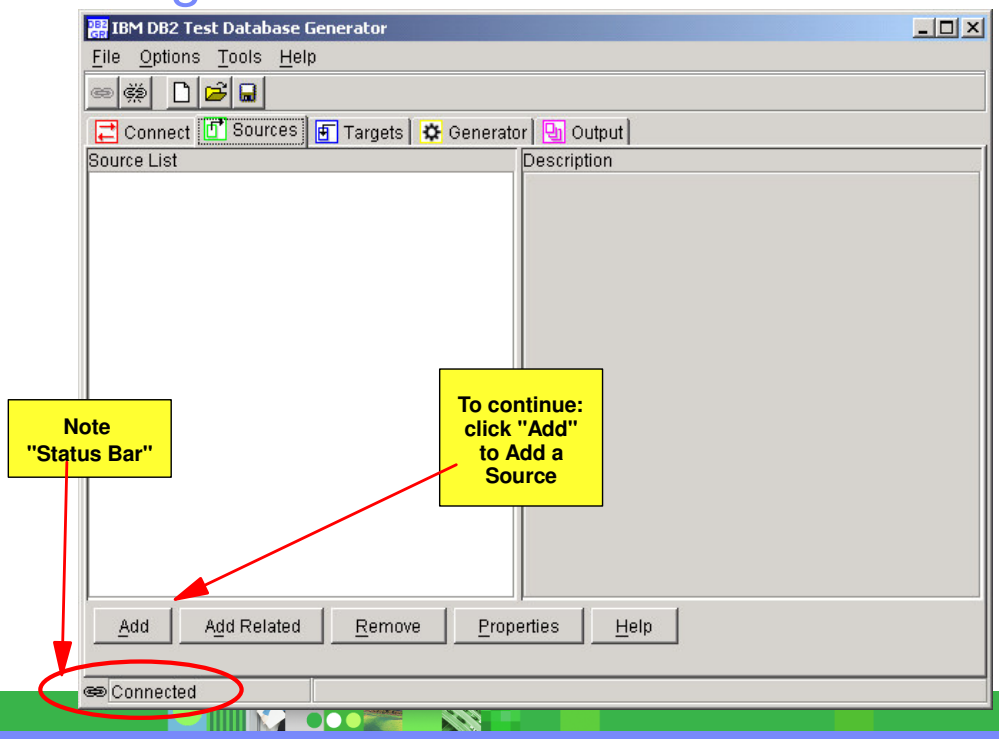


# Initial TDBG ISPF Client Display

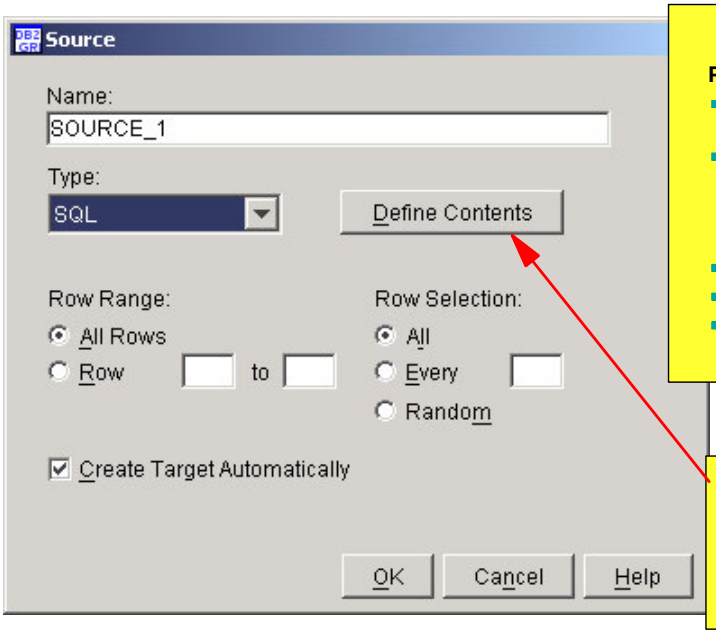





# Adding a Source with Java Client



# Source Type & Row Selection Criteria




**Points of Interest for this display:**

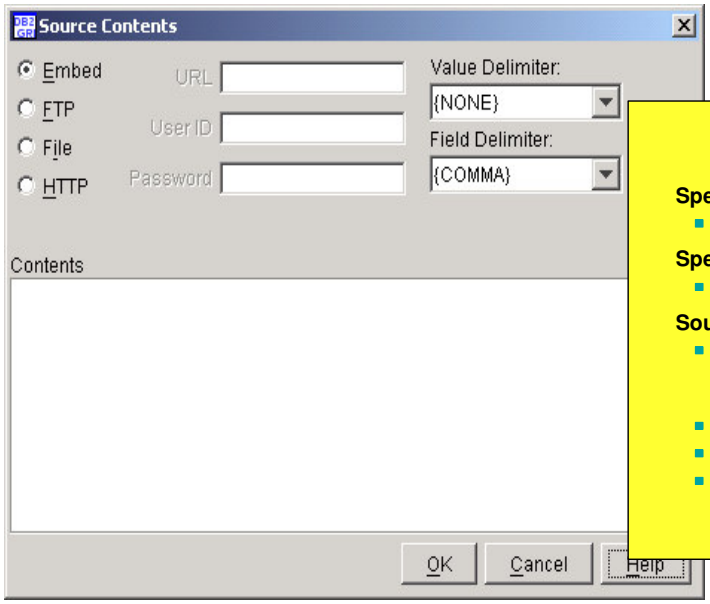
- **Name: Text name you assign to identify the source.**
- **Type:**
  - **SQL - DB2 for z/OS table**
  - **CSV - Delimited Text File**
  - **TXT - Fixed Text File**
- **Row Range**
- **Row Selection**
- **Auto Create option**

**Click "Define Contents" to identify the source object.**

**The window that follows is dependent upon the TYPE specified.**



# CSV Type - Delimited Text File




**Specify Value Delimiter:**

- **none, quote, or apostrophe**

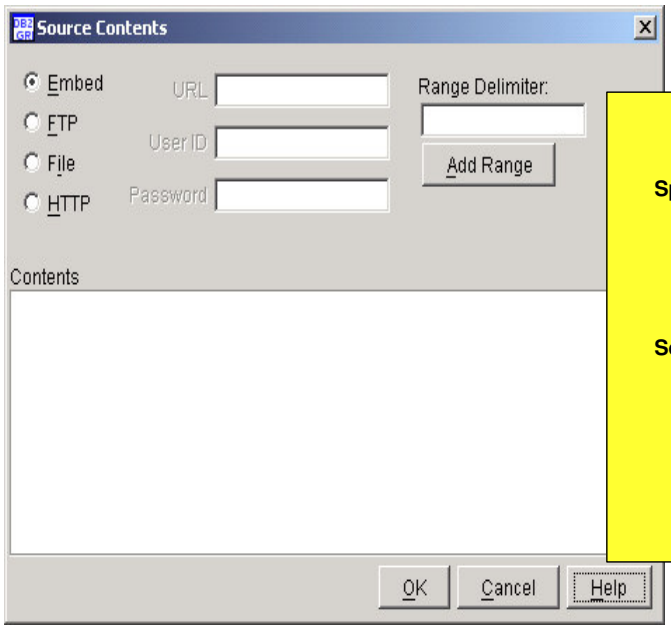
**Specify Field Delimiter:**

- **None, comma, Tab, Semicolon, Space**

**Source contents access methdods:**

- **Embed - data is placed in "Contents" portion of this panel. (not available on ISPF client)**
- **FTP - specify URL, userid, & password**
- **File - specify URL**
- **HTTP - specify URL**

# TXT Type - Fixed Text File




**Specify Value Delimiter:**

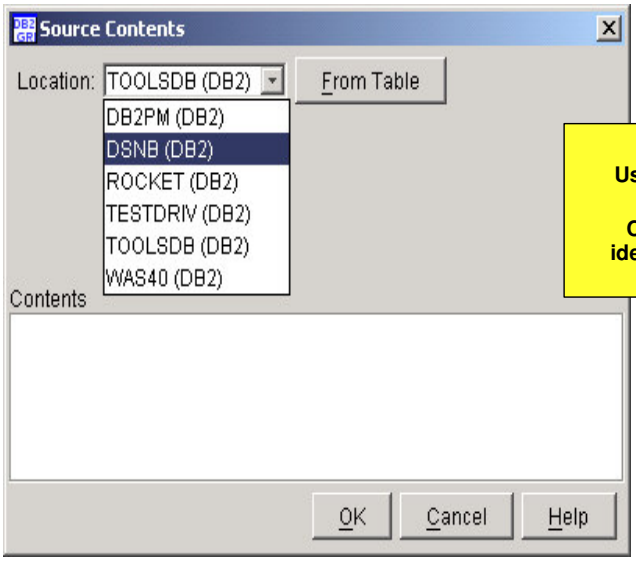
- **Range Delimiter - Comma separated list of "offset" positions within the fixed text file that indicate column positions, must be listed in sequential order. Start at zero.**

**Source contents access methdods:**

- **Embed - data is placed in "Contents" portion of this panel.**
- **FTP - specify URL, userid, & password**
- **File - specify URL**
- **HTTP - specify URL**



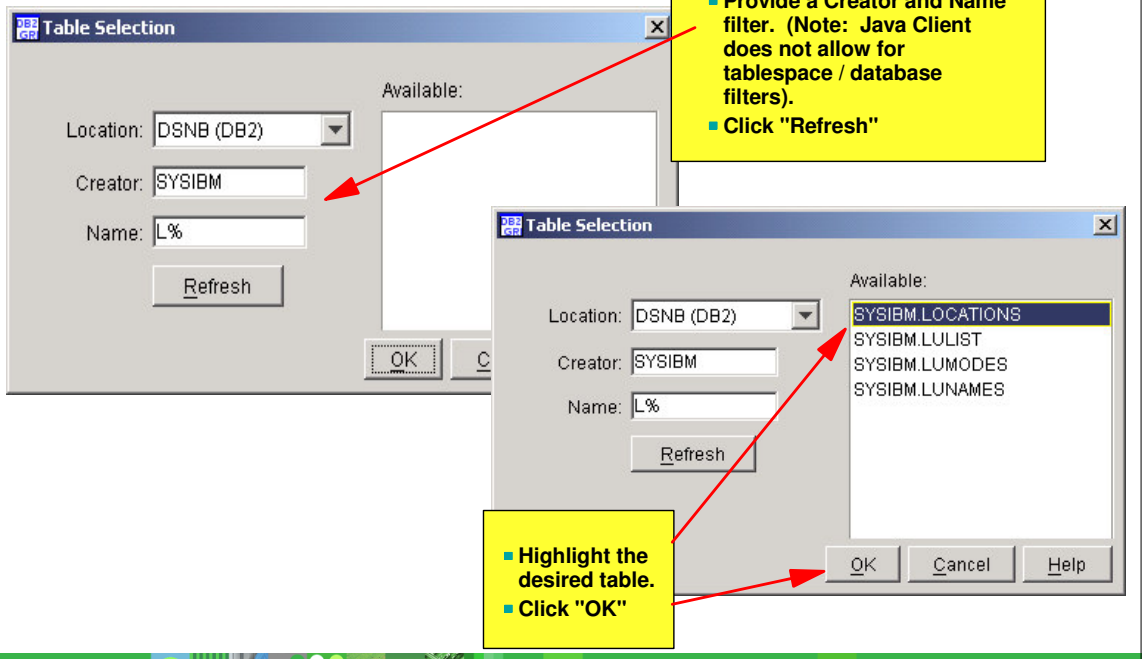
# SQL Type - Location Specification




**Use pulldown to specify source Location.**

**Click "From Table" to identify the source table.**

# SQL Type - Table Identification

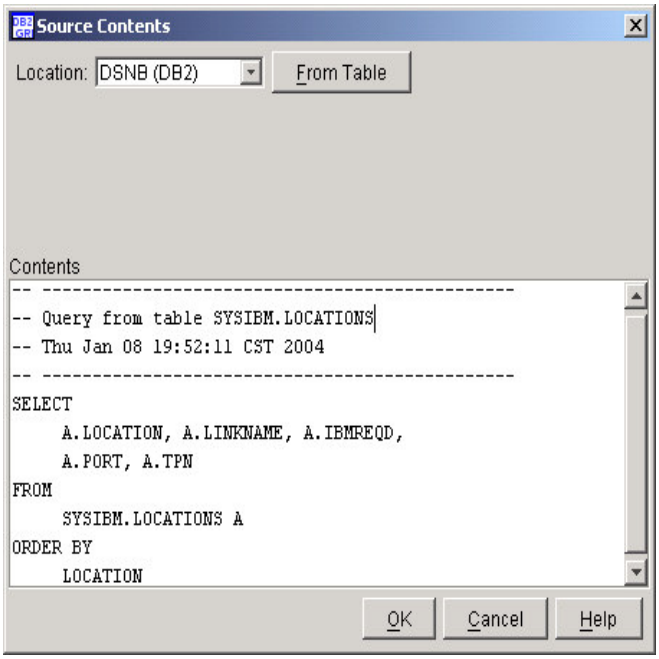



- **Provide a Creator and Name filter. (Note: Java Client does not allow for tablespace / database filters).**
- **Click "Refresh"**

- **Highlight the desired table.**
- **Click "OK"**

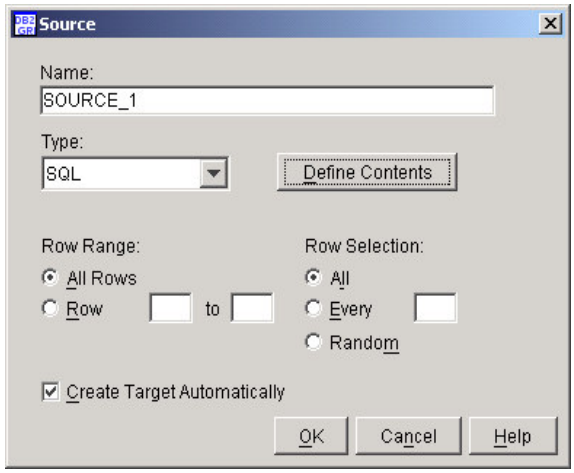


# SQL Type - Editable Query is Displayed




- **If desired, Edit the SELECT statement..**
- **Joins / Unions specified in source SQL specification by either:**
  - **Editing the SQL statement**
  - **Use a View as a source**
- **Click "OK"**

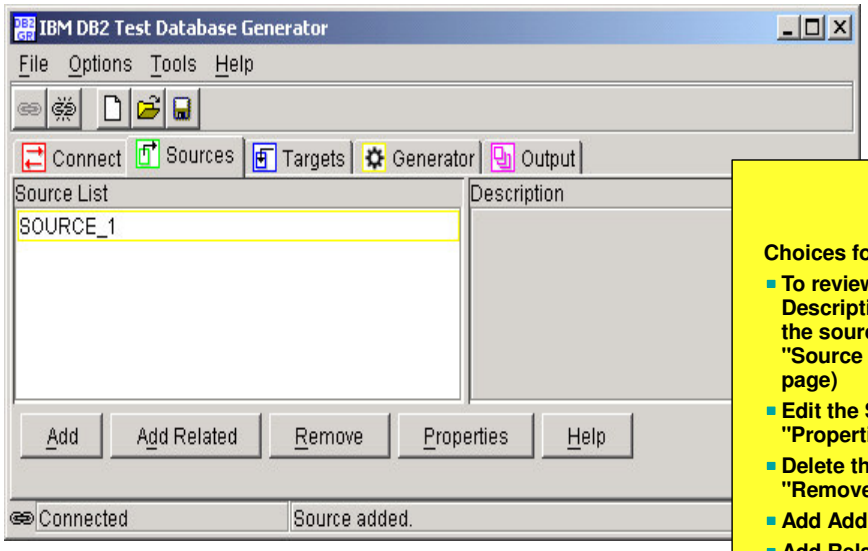
# SQL Type - Source Table Specification Completion




**Click "OK" to return to Source List Window.**



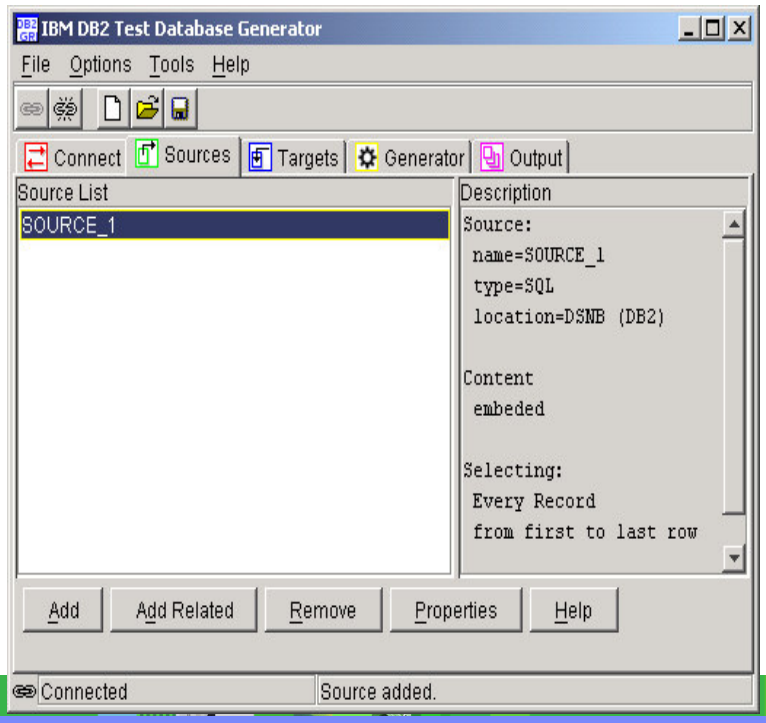
# Source Table Display Summary




**Choices for this display:**

- **To review Source Description, single click the source item in the "Source List". (see next page)**
- **Edit the Source, click "Properties"**
- **Delete the source, "Remove".**
- **Add Additional source(s).**
- **Add Related Objects.**
- **Click "Targets" tab to proceed.**

# Example of Source Description Display






# Adding Related Objects

- DB2 Grouper for z/OS Version 1.1 is shipped with
  - DB2 Test Database Generator for z/OS
  - DB2 Data Archive Expert for z/OS
- Installation of DB2 Grouper is a Separate Process
- Execution of DB2 Grouper is a Separate Process
- The "Add Related" function of the TDBG's ISPF and Java Client REQUIRES DB2 Grouper.
- DB2 Grouper Provides the Following Capabilities:
  - Define non DB2-enforced referential constraints (Application RI)
  - Unit of Work Discovery - Examines DB2 archive log records for objects updated in same UOW.
  - Group Discovery - discovers relationships between tables combining Application RI, DB2 RI, and UOW Discovery information.
  - Enables Editing of Group Composition

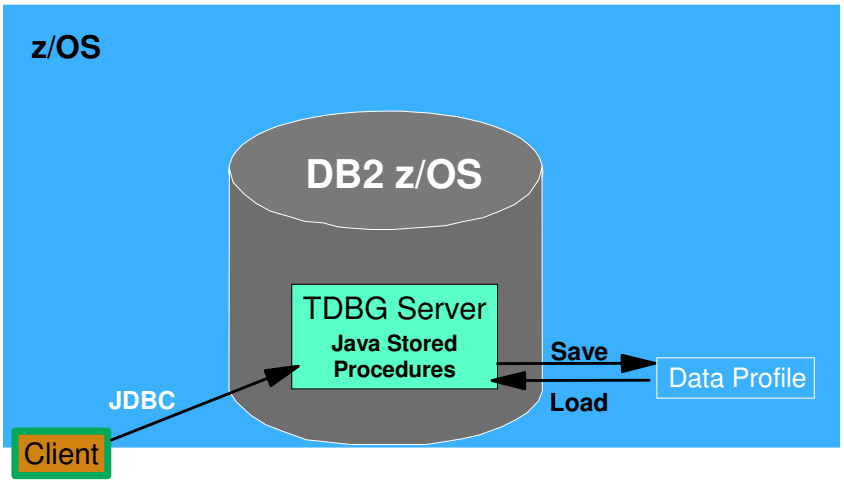
# Source Specification Data Transformation Capabilities (underlined)

- Table
  - Change Table name
  - Change Table Creator
- Columns
  - Add (create) / Exclude
  - Name Change
  - Order of Columns
  - Specify / Change Data Attributes
    - Data Type
    - Length(s)
    - Nullability
  - SQL Transformation
  - Specify Transformation Rule(s):
    - Source Column
    - Static
    - Lookup
    - Mask
    - Expression
    - Random
    - Pattern
- Row Controls
  - Filtering via SQL Predicate
  - Row Range
    - All
    - From x to y
  - Selection within the Range
    - All
    - Every nth Row
    - Random
  - Generate data from scratch
  - Control Number of Rows Generated
- Join / Union Source Tables
- Single Source to Multiple Unique Targets
- Related Tables (Requires Grouper)
  - Identify and Include
- Target DDL Generation

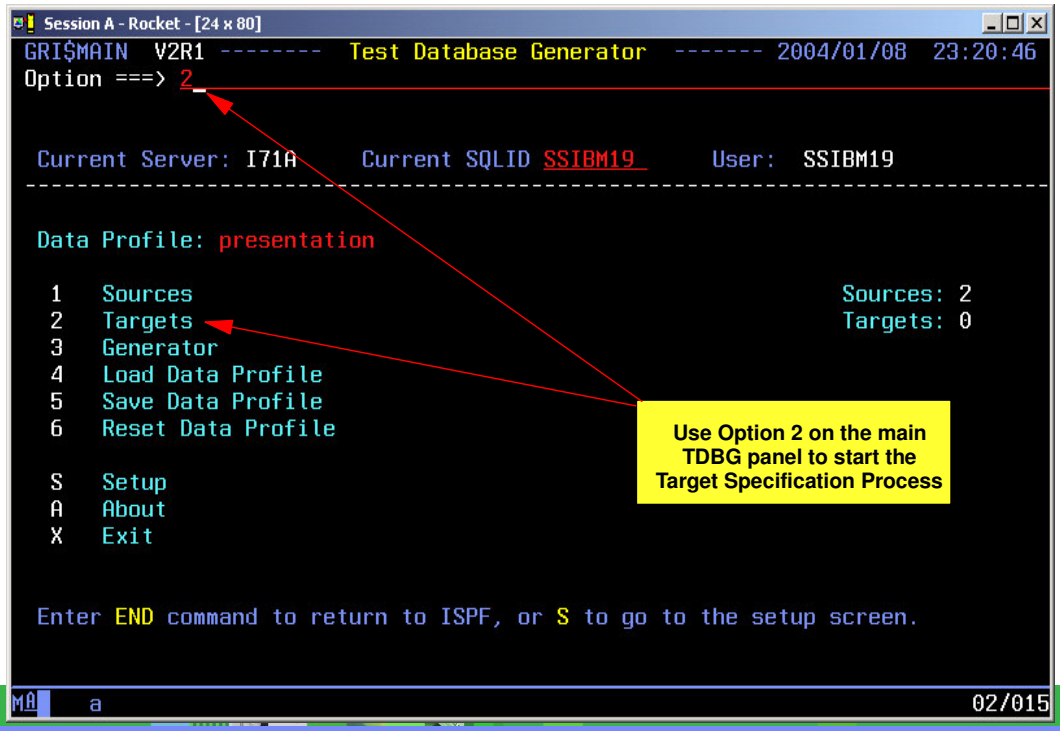


# Target Specification

Source Information kept in the Data Profile is used by Client and user to Build/Define Target Specifications

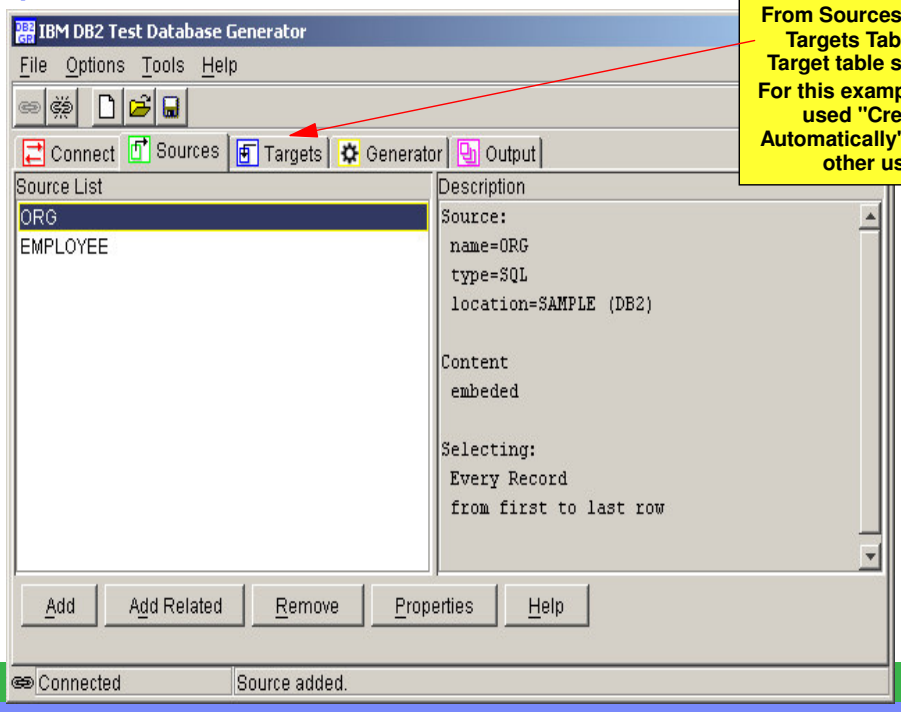


# ISPF Client - Starting the Target Specification

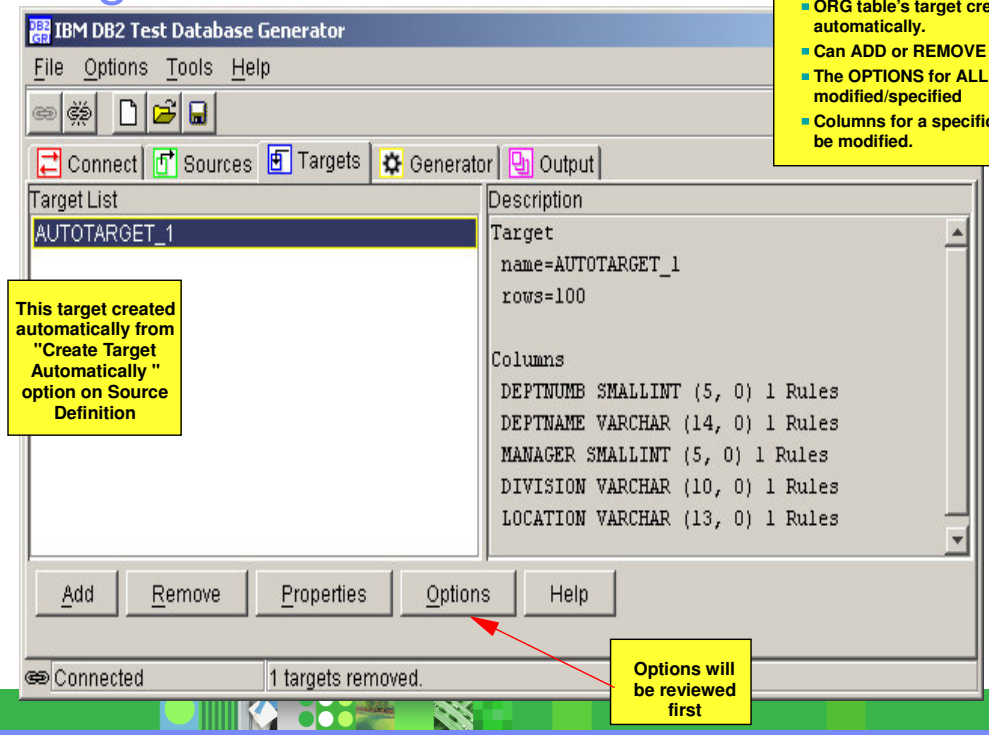





# Java Client - Moving from Source to Target Specification



# Target List



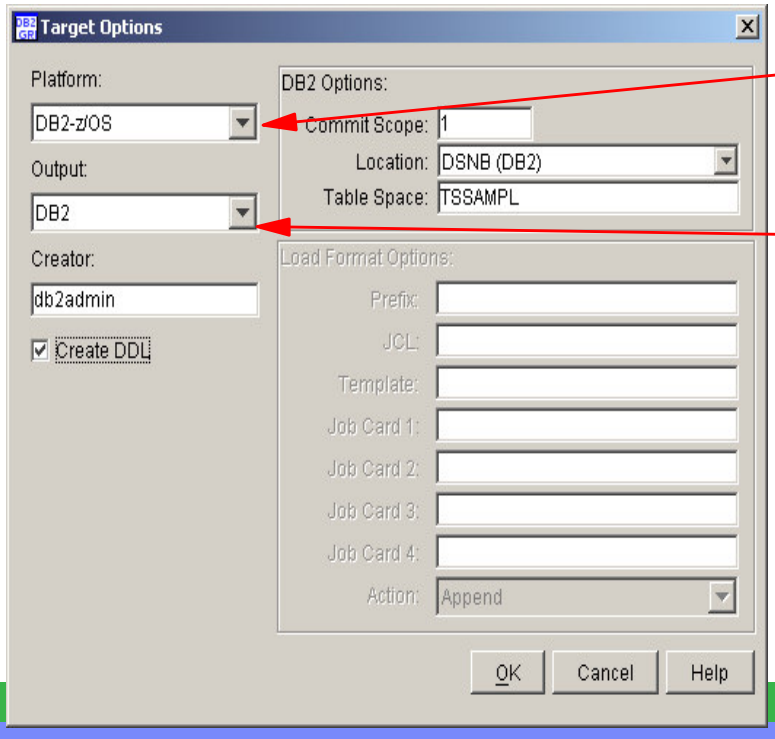

**Points to note:**
- ORG table's target created automatically.
- Can ADD or REMOVE targets
- The OPTIONS for ALL targets can be modified/specified
- Columns for a specific target table can be modified.

**This target created automatically from "Create Target Automatically " option on Source Definition**

**Options will be reviewed first**



# Target Options Window



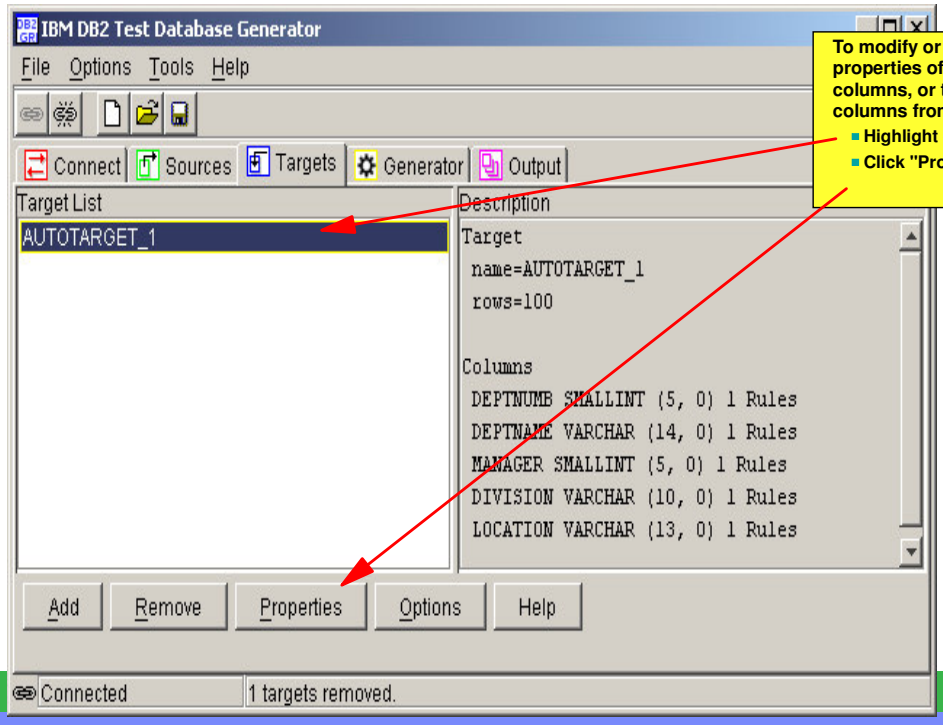

**Target Platform choices:**
- DB2 for z/OS
- DB2 for Multiplatforms

**Output format choices:**
- SQL
- Delimited Text
- Fixed Text
- XML
- DB2 (directly to any)
- DB2 for z/OS Load

**Also Note:**
- Creator id to be used for ALL target tables
- Option to generate DDL for target object.
- Options for DB2 as output type
- Options for DB2 Load output type

**Click OK when completed.**

# Modifying Target Column Properties




**To modify or review the properties of individual columns, or to add/remove columns from the target:**

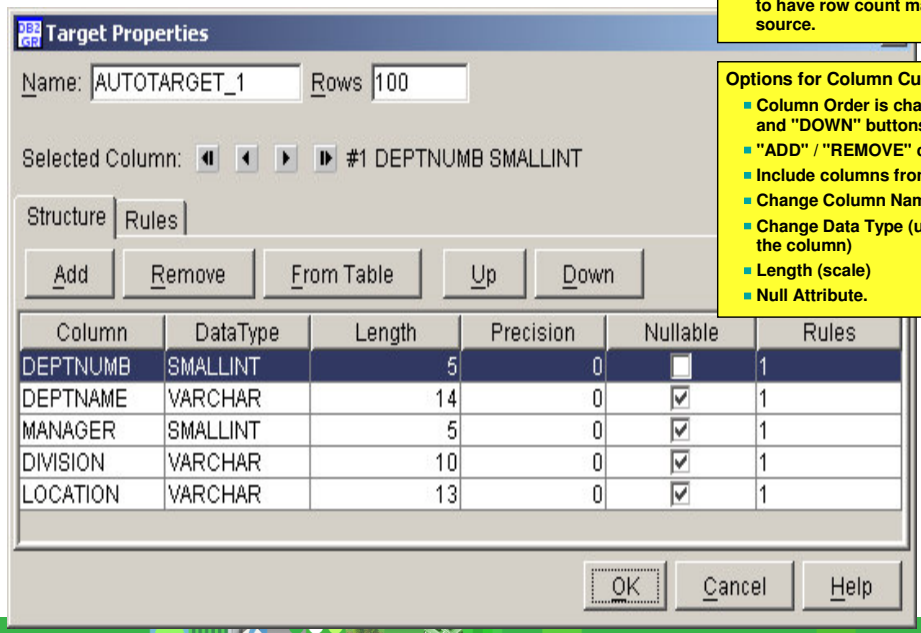
- **Highlight desired table**
- **Click "Properties" button**

IBM Software Group | DB2 Information Management Software

# Column Customization

**Target Object Properties:**

- **Specify Name of target table (Note: creater id specified in Target Options window.**
- **The number of rows inserted WILL MATCH the ROWS specification. Use -1 to have row count match number in source.**

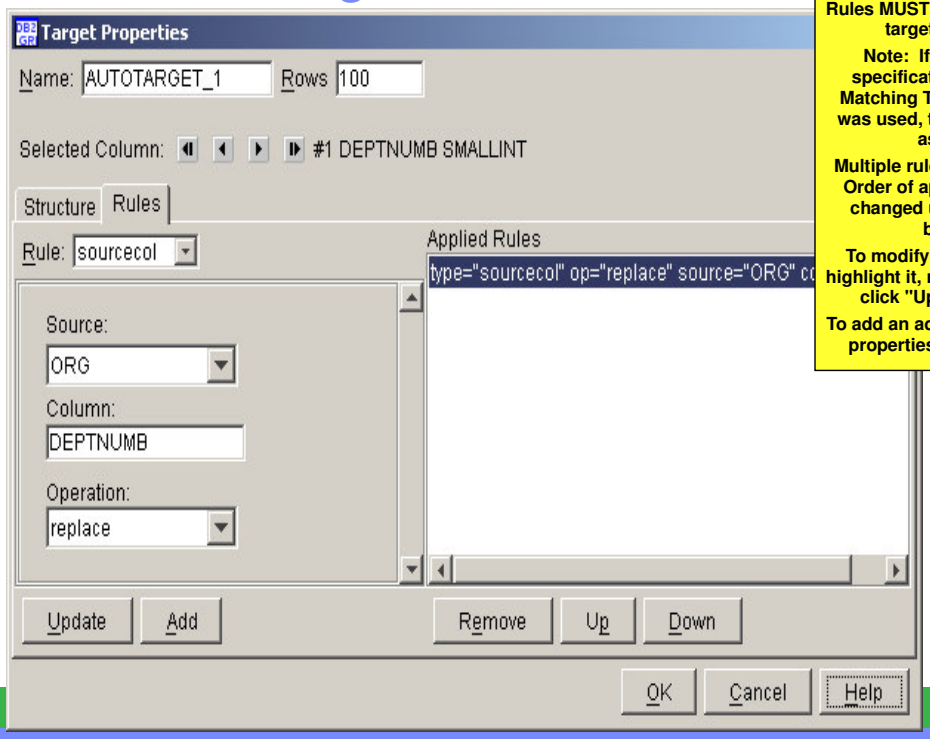



**Options for Column Customization:**

- **Column Order is changed using "UP" and "DOWN" buttons.**
- **"ADD" / "REMOVE" columns with button**
- **Include columns from other tables**
- **Change Column Name by overtyping**
- **Change Data Type (using pulldown on the column)**
- **Length (scale)**
- **Null Attribute.**

**To Review or Modify column RULES, click RULES tab.**

# Rules for Target Columns




**Rules MUST be defined for ALL target columns.**

**Note: If during source specification, the "Create Matching Target = Y" option was used, the "sourcecol" is assigned.**

**Multiple rules can be created. Order of application can be changed using UP/DOWN buttons.**

**To modify an existing rule, highlight it, make changes, and click "Update" button.**

**To add an additional rule, enter properties and click "Add.**

IBM Software Group | DB2 Information Management Software

# Target Specification Data Transformation Capabilities (underlined)

- Table
  - Change Table name
  - Change Table Creator
- Columns
  - Add (create) / Exclude
  - Name Change
  - Order of Columns
  - Specify / Change Data Attributes
    - Data Type
    - Length(s)
    - Nullability
  - SQL Transformation
  - Specify Transformation Rule(s):
    - Source Column
    - Static
    - Lookup
    - Mask
    - Expression
    - Random
    - Pattern
- Row Controls
  - Filtering via SQL Predicate
  - Row Range
    - All
    - From x to y
  - Selection within the Range
    - All
    - Every nth Row
    - Random
  - Generate data from scratch
  - Control Number of Rows Generated
- Join / Union Source Tables
- Single Source to Multiple Unique Targets
- Related Tables (Requires Grouper)
  - Identify and Include
- Target DDL Generation

# Agenda

- Test Data Generation Objectives
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- **Transformations**
  - Source & Target Specifications
  - **Transformation Rules**
- Generation & Results
- Hints & Tips



# Transformation Rules

- Transformation rules define the target test data
  - How to generate test data from source data
  - How to generate test data from scratch
- Examples
  - Create a target column PHONE which is the combination of a country code (derived from COUNTRY file), an area code from TABLE1, and a phone number from TABLE2.
  - Create a target column ACCT_BALANCE which is a random number that falls within a specified range.
  - Create a target column that is exactly the PIN column with the 3rd and 5th positions replaced (masked) with the letter X.

# Rule Terminology - Scopes and Sets

- You define your test data one target column at a time
- The scope of a transformation rule set is target column
- Multiple rules can be specified for each target column
- Transformation rules are applied in order
- Each rule can modify, replace, append, or preface the previous value to allow for incremental building of a target column

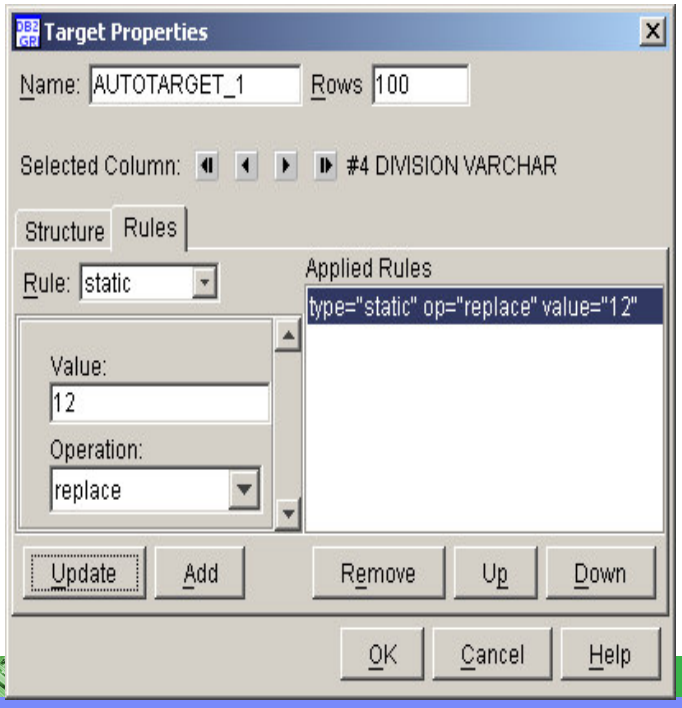


# Transformation Rules

- Static Values
- Source Column Values
- Data Lookup
- Data Masking
- Expressions
- Random Values
- Pattern Generation

# Rule 1: Static Values

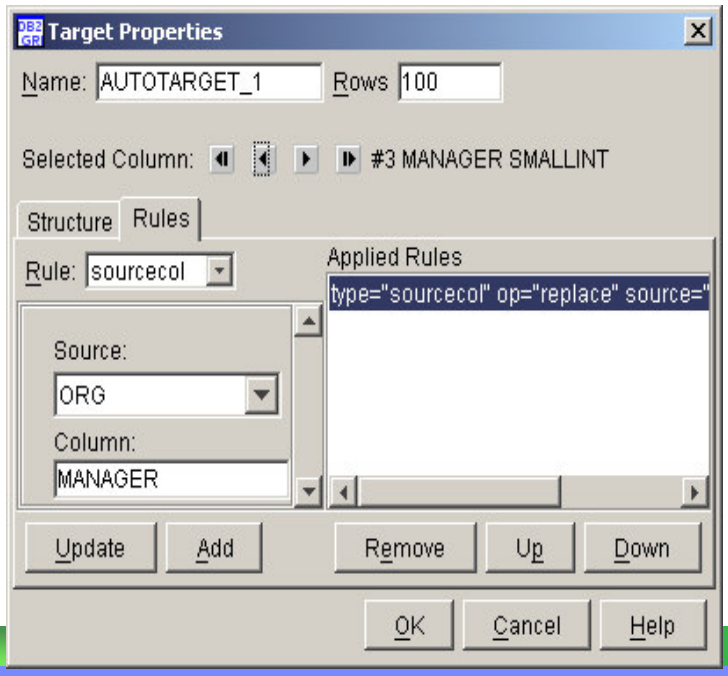
- Specify static data value
- Source data not used as input for this rule
- Examples
  - Set target column DIVISION to 12
  - Set target column STATE to CA
  - Set target column EXP_DATE to 2030-12-31




IBM Software Group | DB2 Information Management Software

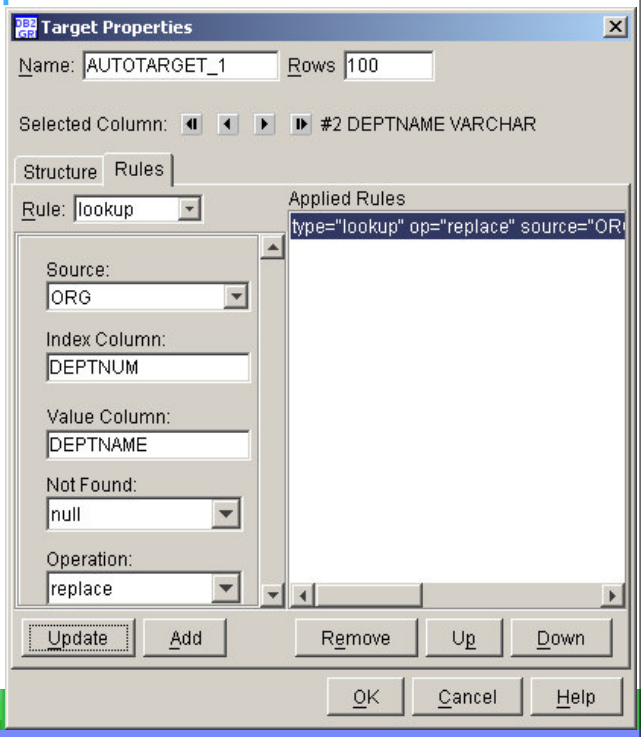
# Rule 2: Source Column Values

- Generate target column value based on source column value
- No transformation (use source data as-is)
- Copies a column value from the currently selected row in a source object
- Used to perform a “copy” function
- Used to supply initial column value from a data source
- TDBG Auto-Target feature uses this rule as the default



# Rule 3: Data Lookup

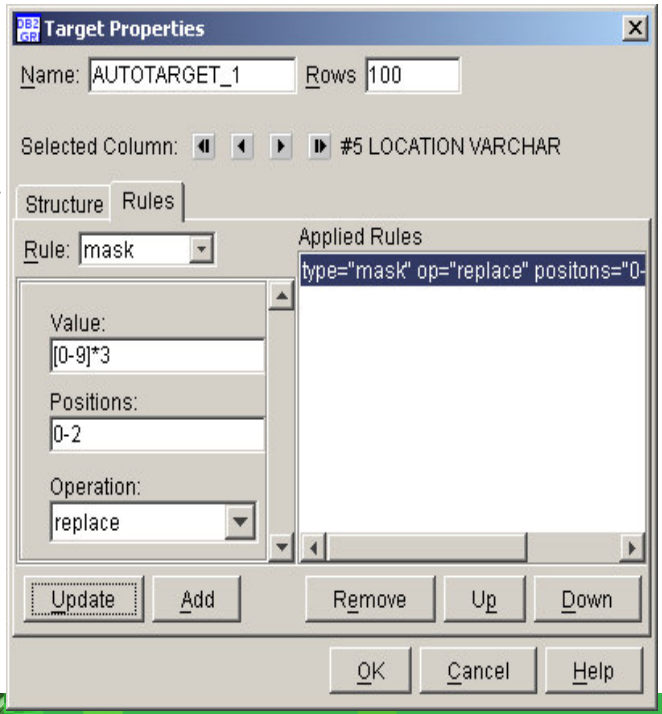
- Provides a method to replace data based on table lookups
- Uses the current value of the generated column as a key to the lookup
- Specify replacement column
- Example:
  - Source data has a product code which needs to be represented as a product name
    - PROD_NAME = LOOKUP(PRODUCT_CODE in PRODUCT_TABLE)






# Rule 4: Mask

- Replace or hide sensitive data
- Masks can be set with static text
  - Replace account number with a string of X's
  - Replace last four digits of License ID with 9999
- Mask can be set using a pattern rule (see rule #7)
  - Replace first character with a letter in the range A-F and then replace the next position with a number between 0 and 9
- Mask can be applied to the entire column or to a substring of the column

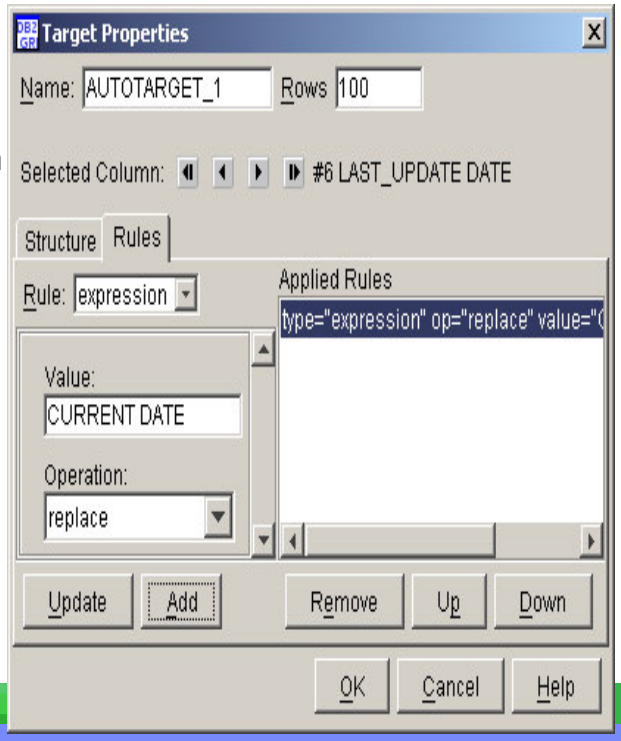


# Rule 5: Expression

Expression provides a method to call database specific functions

- Supports any expression that can be evaluated by the DBMS in which TDBG is installed
  - String manipulation, calculations, etc.
- Examples
  - Calculate the current date / time / timestamp
  - Evaluate a mathematical expression
    - Target column =

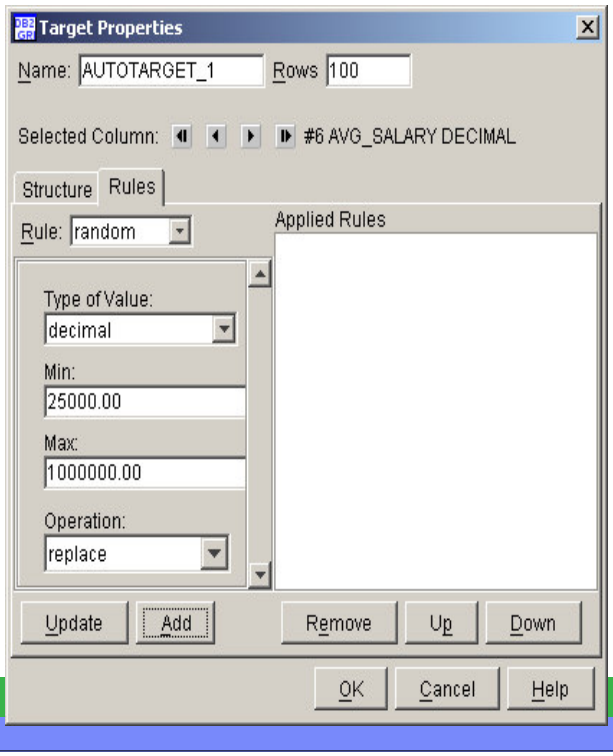
      source column * 1.1





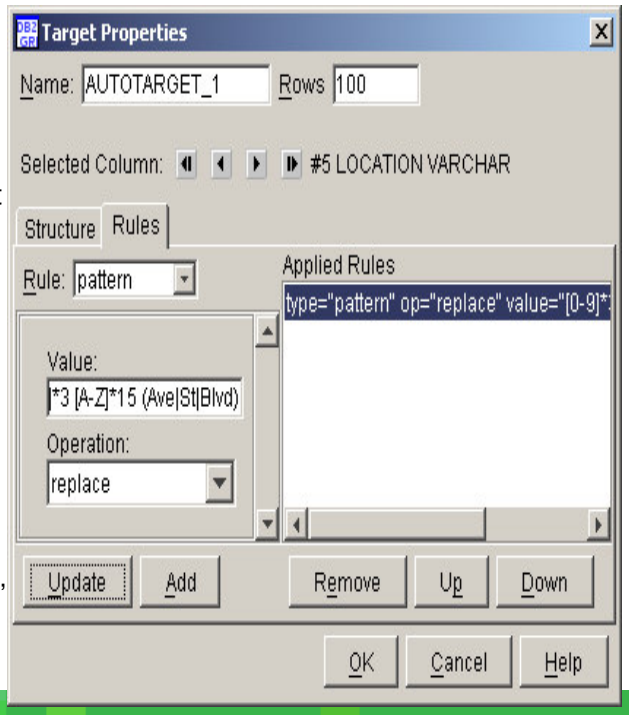

# Rule 6: Random

- Generate a random value
- Allows for creating random date, time, timestamp, integer, and decimal values
- Specify min / max ranges for the generated values
- Randomly generated values are propagated across related tables
  - Only applies to primary/foreign key relationships that are system or user defined
  - A DEPT_CODE that is randomly generated shows up in both the DEPT table and the EMPLOYEE table



# Rule 7: Pattern

- Generate data based on a specified pattern
- Numeric pattern
  - [0-9] evaluates to any single-digit number
- Character pattern
  - [A-Z]*3 evaluates to any three-character uppercase string
  - Character patterns are randomly selected at generation runtime
- String pattern
  - (Mrs|Mr|Ms) evaluates to 'Mrs', 'Mr', or 'Ms'.
  - (C[ATO] | A[KLR]) evaluates to 'CA', 'CT', 'CO', 'AK', 'AL', or 'AR'



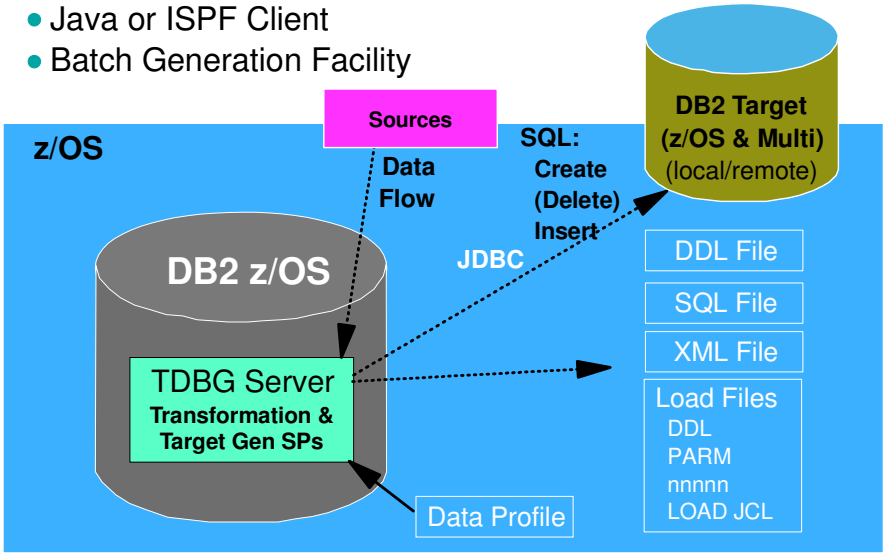

IBM Software Group | DB2 Information Management Software

# Agenda

- Test Database Generation Objective
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- **Generation & Results**
- Hints & Tips

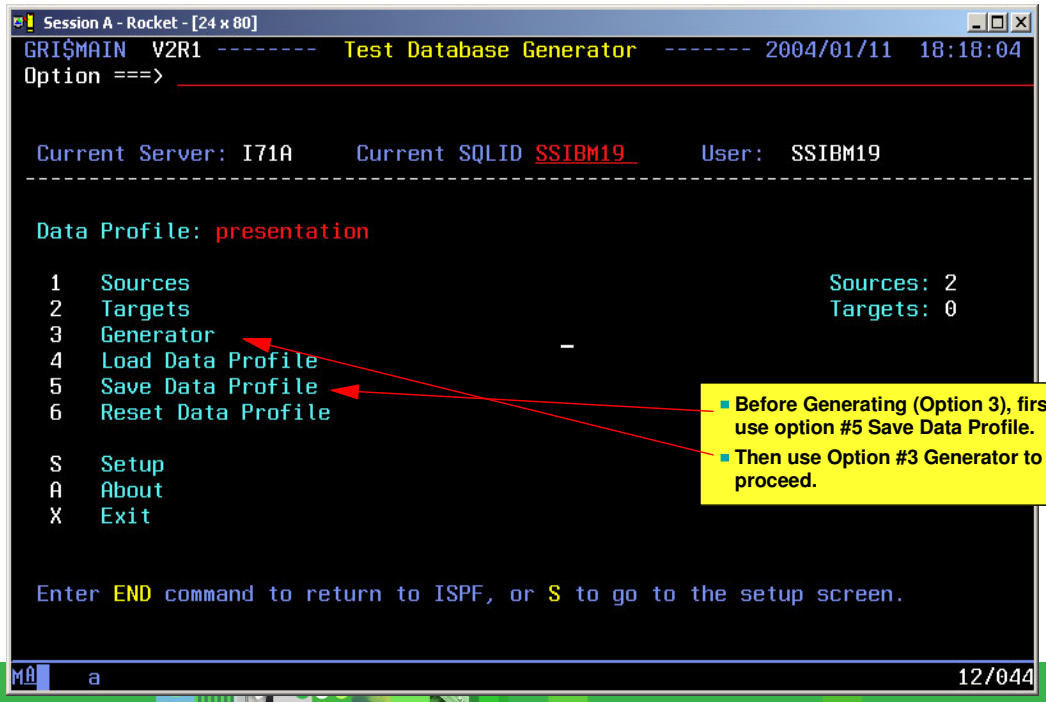
# TDBG Generation

- Initiated by:
  - Java or ISPF Client
  - Batch Generation Facility

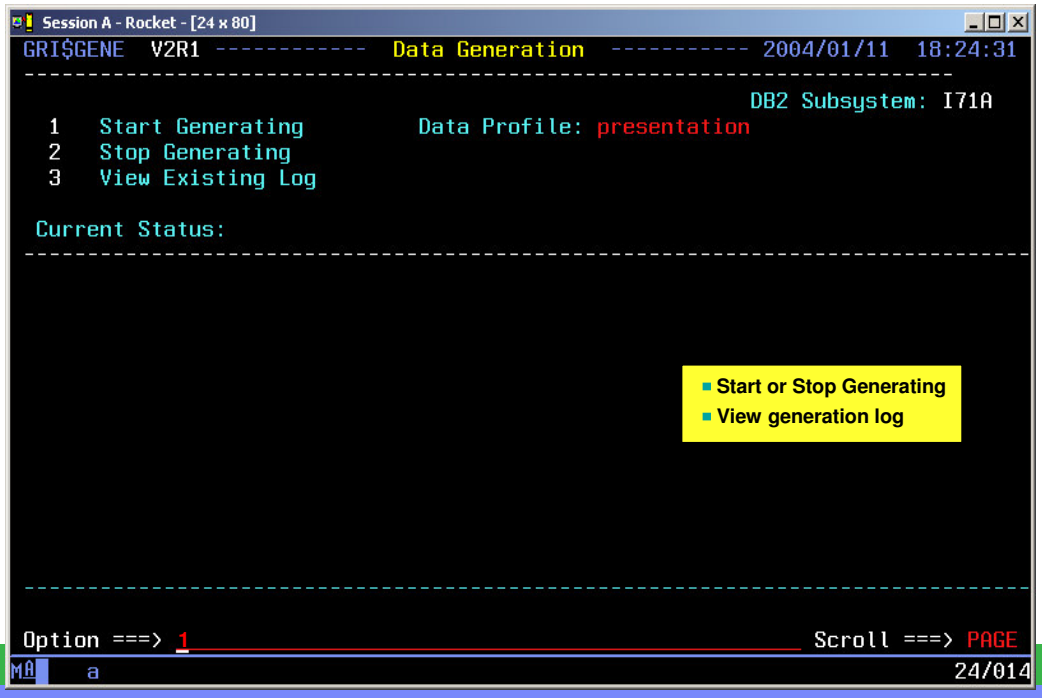





# Data Generation



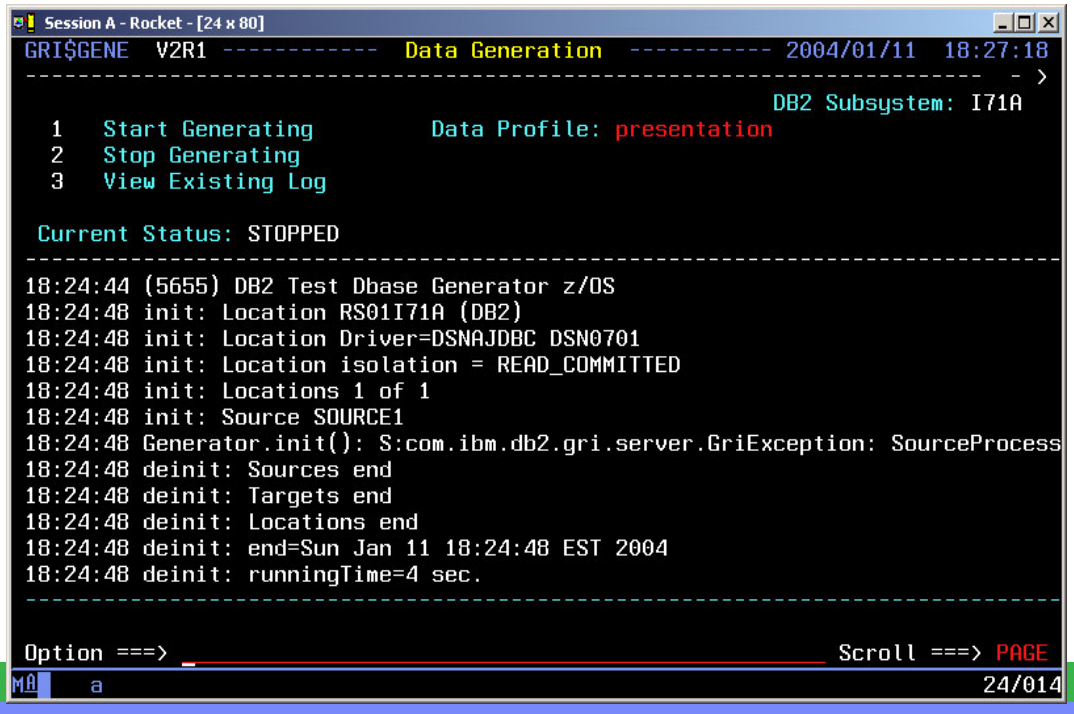
# Start the Generation Process




- Start or Stop Generating
- View generation log



# Generation Results Log



# Agenda

- Test Database Generation Objective
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- **Hints & Tips**

TDBG = Test Database Generator



# Usage Hints - ISPF Panel Display Customization

- Some ISPF panel fields contain long fields and are left/right scrollable
- TDBG supports Long Object Names, default ISPF Client displays are set to max column width
- Some users Want to Control Other Panel Display Options
  - CSETUP command when executed on ISPF panels with column objects enables customization of following ISPF columns attributes:
    - FIX - keep selected columns on Left Side of Panel
    - ORDER - control horizontal placement of columns
    - SIZE - size of the column
    - SORT - select column(s) to sort
    - RESET - rollback recent customizations
    - REMOVE - use product defaults
    - Make Customization Permanent or Temporary for User doing customization
  - See Appendix B of TDBG User's Guide

# Usage Hints - Downloading the Java Client

- Must be FTPed from Host to Workstation. In addition, documentation does not identify the proper file to download.
- FTP in binary
- Instructions for Locating and FTPing Java Client code from Windows workstation:
  - Start a DOS prompt window
  - cd to the directory that you will place the code: e.g. c:\temp\tdbg
  - ftp
  - open <hostname> (e.g. Dallas demo hostname: demomvs.demopkg.ibm.com
  - enter host userid when prompted
  - enter host password when prompted
  - cd /usr/lpp/griv2r1/client (or other directory if default install dir not used)
  - binary
  - get TDBGClient32.exe (file is about 40 meg)
  - quit
- Using the "Run" window, locate the file and execute it.



# Additional Hints

- RULE descriptions and examples are documented in Chapter 3 and Appendix D of the DB2 Test Database Generator User's Guide.
- Installation and Configuration requires following skill sets in addition to typical skills used in installing z/OS systems software products:
  - z/OS
    - z/OS ISPF
    - Java Virtual Machines (JVM)
    - Working with DB2 for z/OS and JDBC installation & configuration
    - Working with OMVS segments
    - Unix Systems Services (USS)
    - DB2 for z/OS Stored Procedure Address Space Enablement and using SQL to create stored procedures.
    - Workload Manager (WLM)
  - Windows
    - DB2 Connect Installation and Configuration
- DB2 Grouper provides ALL Referential Integrity functionality.
- Any users of DB2 TDBG must have a complete OMVS segment with a shell and home directory

# Documentation & Reference Material

| Document Name | Document Number |
|---|---|
| Program Directory for DB2 Test Database Generator for z/OS Version 2.1 | GI10-8516 |
| IBM DB2 Test Database Generator User's Guide | SC18-7411 |
| Fact Sheet - TDBG | GC18-9148 |
| Program Directory for DB2 Grouper for z/OS Version 1.1 | GI10-8569 |
| IBM DB2 Grouper User's Guide | SC18-7409 |

Website for all DB2 Tools, including links to product information, documentation and support information:
http://www.software.ibm.com/data/db2imstools



# Agenda

- Test Database Generation Objective
- IBM DB2 Test Database Generator for z/OS Version 2.1 Today
  - Summary of Capabilities
- Terminology and Architecture
- Potential Configurations
- Transformations
  - Source & Target Specifications
  - Transformation Rules
- Generation & Results
- Hints & Tips