

# PureData Systems for Analytics (Netezza) Connectivity for InfoSphere Information Server DataStage



IBM

## Agenda

- Download Netezza Client
- Basic Netezza Commands
- Configuring DataStage for Netezza
  - Netezza Connector (Available at 8.5 Fix Pack 1 and later)
  - Netezza Enterprise Stage (Deprecated)
    - Connector Migration Tool
- Balanced Optimization with Netezza Connector

a stnarter

- Troubleshooting
  - Debug Environment Variables
  - Netezza Logs
  - ODBC Driver Manager Tracing
  - ODBC Driver Tracing



## Download Netezza Client

- Netezza is an appliance
- Netezza Emulator is available for development
  - IBM PureData-Netezza Developer Network
- Netezza Client Software
  - IBM Fix Central
- Unix Installation:
  - Once downloaded from Fix Central and extracted, several archive packages will exist with an unpack script
  - Run the unpack script to install the Client/ODBC drivers to a specified directory.
  - Leverage the DataDirect Driver Manager shipped in /IBM/InformationServer/Server/branded\_odbc to manage the new Netezza ODBC driver
- Windows Installation:
  - Run the appropriate setup program(s)
  - nzodbcsetup.exe (required for Netezza Connector, Netezza Enterprise Stage, ODBC Connector)
    - Ensure to install 32 bit drivers Information Server DataStage 11.3 and earlier Engine is 32 bit on Windows
  - nzsetup.exe (required for nzload in the Netezza Enterprise Stage)

tor a smarter plane



### **Basic Netezza Commands**

How to \ Database	Command
Start System	Run nzstart. Example: From Netezza Server: /nz/kit/bin ./nzstart
Connect to SQL Command Interpreter: nzsql	Set Environment Variables: NZ_USER=admin NZ_DATABASE=system NZ_PASSWORD=password Encrypt password using ./nzpassword to create a locally stored encrypted password. Run nzsql. Example: From Netezza Server: /nz/kit/bin [nz@netezza bin]\$ nzsql SYSTEM(ADMIN)=>
Connect from a client using: nzodbcsql	Run nzodbcsql. Example: From Client: /opt/Netezza/bin ./nzodbcsql -h netezza-slb -d support -u team2 -pw team2
Stop System	Run nzstop. Example: From Netezza Server: /nz/kit/bin _/nzstop



# for a smarter planet M **Basic nzsql Commands**

How to \ Database	Command
Create database	From nzsql prompt run: SYSTEM.ADMIN(ADMIN)=>CREATE DATABASE SUPPORT;
Drop database	From nzsql prompt run: SYSTEM.ADMIN(ADMIN)=>DROP DATABASE SUPPORT;
Connect to database	From nzsql prompt run: SYSTEM.ADMIN(ADMIN)=>\c DATABASE USERNAME PASSWORD Example: SYSTEM.ADMIN(ADMIN)=> \c support team2 team2
Show Databases	From nzsqlpl prompt run: SYSTEM.ADMIN(ADMIN)=> \I
Show tables	From nzsql prompt run: SYSTEM.ADMIN(ADMIN)=>\dt
Describe Table	From nzsql prompt run: => \d TABLE_NAME
Grant Roles	For Details see Information Center
Enter a query	Type the query directly in the nzsql prompt Example: SUPPORT.ADMIN(TEAM2)=> select * from MyFirstTable;



# for a smarter planet M **Additional Commands**

How to \ Database	Command
Show current sessions	From nzsql run: => \act
Show version	From Netezza Server: /nz/kit/bin ./nzrev

emart



## Configuring DataStage for Netezza

- Three ways to connect to Netezza using DataStage:
  - Netezza Connector (Preferred method for all new development)
  - Netezza Enterprise Stage (Deprecated)
  - ODBC Connector
- All connection methods leverage ODBC connections
- Netezza Connector
  - Unloads and Loads are done by using ET(External Tables) and loads also use TWT (Temporary Work Tables)
  - Link to Information Center
- Netezza Enterprise Stage (Deprecated)
  - Unloads and Loads can by done by using ET (External Tables) Updates/Deletions also use TWT
  - The Netezza Enterprise stage allows for Netezza Loads to be done with the nzload command
  - Link to Information Center
- ODBC Connector
  - Pure ODBC connection



### Windows ODBC Driver Manager

Ensure to configure the DSN in the Windows 32 bit ODBC Driver Manager

Located here on 64 bit Windows installs:

#### C:\Windows\SysWOW64\odbcad32.exe

4	ODBC Data Source Administrator (32-bit)	×
User	IBM Netezza ODBC Driver Setup	
Sys	DSN Options Advanced DSN Options SSL DSN Options Driver Options	
N	Data Source: NZSQL Description: Scott's Netezza Emulator	
	Server: netezza-slb.uspa.ibm. Port: 5480	
	Database: support	
	User Name: sbrokaw Password: ••••••	I
	Test <u>C</u> onnection	
	OK Cancel Help	
	OK Cancel Apply Help	

## Sample ODBC Entry (Unix)

[netezza-slb]	
Driver	= /opt/Netezza/lib64/libnzodbc.so
Description	= Scott's Netezza Emulator
Servername	= netezza-slb.uspa.ibm.com
Port	= 5480
Database	= support
Username	=
Password	=
ReadOnly	= false
ShowSystemTables	= false
LegacySQLTables	= false
LoginTimeout	= 0
QueryTimeout	= 0
DateFormat	= 1
NumericAsChar	= false
SQLBitOneZero	= false
StripCRLF	= false
securityLevel	= preferredUnSecured
caCertFile	=

Similar entry for each platform can be found in the Netezza Client install directory: /NetezzaClient/lib/odbc.ini.sample

Ensure to add the DSN to the top entry: [ODBC Data Sources]

netezza-slb=Scott's Netezza Emulator



## Configuring DataStage for Netezza

- (Unix Only) Netezza ODBC driver expects odbc.ini instead of .odbc.ini
- (Unix Only) Create Symbolic link: cd /IBM/InformationServer/Server/DSEngine In -s .odbc.ini odbc.ini
- Add Environment variables to dsenv file or Windows System Variables (for Netezza Connector and Netezza Enterprise Stages only)
- Restart ASBAgents, DataStage Engine after making environment changes

### IBM

## Netezza Environment variables

Name	Required for Stage	Description
NETEZZA	Netezza Enterprise	The location of the Netezza Client installation. Example: /opt/Netezza
LIBPATH or LD_LIBRARY_PATH or SHLIB_PATH	Netezza Enterprise (Unix Only)	The location of the Netezza client libraries Example: \$NETEZZA/lib
PATH	Netezza Enterprise	The location of the Netezza client executables. Example: \$NETEZZA/bin
ODBCINI	Netezza Enterprise Netezza Connector	ODBCINI should be set in the dsenv file and points to the path to the .odbc.ini file Example: /opt/IBM/InformationServer/Server/DSEngine/.odbc.ini This variable is already set in default desenv files
NZ_ODBC_INI_PATH	Netezza Connector	The location of the odbc.ini file. NOTE: This is the odbc.ini NOT the .odbc.ini file – although the locations of the files can be the same – i.e. both in \$DSHOME Example: /opt/IBM/InformationServer/Server/DSEngine

a smarter planet M



### Configuring nzload on 64 bit Linux

- nzload is a 32 bit application
- On 64 bit Linux Systems, nzload has dependencies on Linux 32 bit libraries

for a smarter planet

- SSL Libraries
  - libssl
  - libcrypto
- See Technote <u>1470676</u> for assistance in resolving
  - Install the 32-bit version of SSL and CRYPTO libraries
  - Make sure the binaries have the required permissions

```
sh-4.1$ ldd `which nzload`
        linux-gate.so.1 \Rightarrow (0x00a24000)
       libnzodbc nzload.so => /opt/Netezza32/bin/libnzodbc nzload.so (0xf64aa000)
       libm.so.6 => /lib/libm.so.6 (0x005c0000)
       libc.so.6 => /lib/libc.so.6 (0x001a0000)
       libpthread.so.0 => /lib/libpthread.so.0 (0x00ee3000)
       libdl.so.2 => /lib/libdl.so.2 (0x00a83000)
       libssl.so.6 => /usr/lib/libssl.so.6 (0x00337000)
       libcrypto.so.6 => /usr/lib/libcrypto.so.6 (0x00d2a000)
        /lib/ld-linux.so.2 (0x0017e000)
       libgssapi krb5.so.2 => /lib/libgssapi krb5.so.2 (0x00915000)
       libkrb5.so.3 => /lib/libkrb5.so.3 (0x00b84000)
       libcom err.so.2 => /lib/libcom err.so.2 (0x00ce9000)
       libk5crypto.so.3 => /lib/libk5crypto.so.3 (0x00543000)
       libresolv.so.2 => /lib/libresolv.so.2 (0x00cfd000)
       libz.so.1 => /lib/libz.so.1 (0x00382000)
        libkrb5support.so.0 => /lib/libkrb5support.so.0 (0x00396000)
        libkeyutils.so.1 => /lib/libkeyutils.so.1 (0x00f8e000)
        libselinux.so.1 => /lib/libselinux.so.1 (0x008d3000)
sh-4.1$
```



## Configuring DataStage for Netezza

- When designing a parallel Information Server DataStage job, the data source names are not appearing in the Netezza connector when clicking on the Select button
- See Technote <u>1628099</u> for assistance resolving
  - Ensure that there are no spaces on either side of the = symbol
  - Ensure that word "Netezza" appears in the description of the DSN

[ODBC Data Sources] DB2 Wire Protocol=D  netezza-slb = Scott	ataDirect DB2 Wire Protocol Driver 's Netezza Emulator	<pre>[ODBC Data Sources] DB2 Wire Protocol=DataDirect DB2 Wire Protocol Driver   netezza-slb=Scott's Netezza Emulator</pre>	
General Properties Columns Advanced	Partitioning	General Properties Columns Advanced Partitioning	
Data source *		✓ Connection Test Load Sar	ve 🔺
Database *	DataStage	Data source *	
Usemame * Password *		Databa Usema	
Use separate connection for TWT	There are no values available for this property.	Passwo Data source isSustem description	
▼ Usage		▶ Use netezza-slb true Scott's Netezza VM	
Write mode		▼ Usage Dat	<u>a</u>
Table name *	ок	Write m	
Enable case-sensitive identifiers			

# Information Management of Ware for a sharter planet

## Configuring DataStage for Netezza Connector – Importing metadata

° ai	nd QualityStage	Designer			
У	Diagram Debug	Import Export Tools Window Help	_		
	🛱 🔗 🐼 🔎	DataStage <u>C</u> omponents	1 🐨 🥖 🕅 🖨 🖨 🎑 🚱 🗌		
	Devellel	DataStage Components (XML)			
FI	Parallel	External Eunction Definitions			
H 11		Schema Library Manager			
H		Web Service Function Definitions			
H		Via <u>B</u> ridges		Connector metadata import	×
H		Table Definitions	Start Conn <u>e</u> ctor Import Wizard		
		IMS Definitions	Assembler File Definitions		
+			<u>C</u> OBOL File Definitions	and the second free level of	
Ľ			DCLGen File Definitions	pr to be used for import.	

#### Connector selection

#### Connectors:

Name	Туре	Variant	Hosted on	
DB2 Connector	DB2Connector	9.1	HARVEY.	SWG.US
Greenplum Connector	GreenplumConnector	4.2	HARVEY.S	SWG.US
Netezza Connector	NetezzaConnector	4.5	HARVEY.	SWG.US
ODBC Connector	ODBCConnector	3.5	HARVEY.	SWG.US
Oracle Connector	OracleConnector	10	HARVEY.S	SWG.US
Oracle Connector	OracleConnector	11	HARVEY.	SWG.US
Teradata Connector	TeradataConnector	12	HARVEY.S	SWG.US
		_	1	



#### Netezza Connector – Default Options Netezza\_Connector\_Target - Netezza Connector Stage Input N - 1 Netezza\_Connector\_Source - Netezza Connector Input name (upstream stage) Stage Output NZ\_Output (Row\_Generator\_4) Output name (downstream stage) 🦯 NZ\_Input () Ŧ General Properties Columns Advanced Partitioning Connection Test Load Save Visage General Properties Columns Advanced Write mode Insert Test Load Save Connection Table name MyTable Enable case-sensitive identifiers No NZSQL Data source Truncate column names No Database ' support SQL Usemame team2 Direct insert No ----Password Atomic mode Yes Enable record ordering No Generate SQL at runtime Yes Key columns \* MyTable 2 Use unique key column No Table name \* Enable case-sensitive identifiers No Action column \* SQL Check duplicate rows No Select statement \* User-defined SQL \* Enable partitioned reads No Table action \* Append Transaction Session Record count 2000 Schema reconciliation Mark end of wave No Unmatched link column action Drop Type mismatch action Drop Session Unmatched table column action Ignore nullable Schema reconciliation Mismatch reporting action Warning Unmatched link column action Drop Temporary work table mode Automatic Type mismatch action Drop Table name \* Mismatch reporting action Warning Truncate table No Unload options Drop table Yes Load options Directory for named pipe (Unix only) Generate statistics No Before/After SQL No Maximum reject count 1 Limit number of returned rows No Directory for named pipe (Unix only) Directory for log files Other options Before/After SQL No

tware for a smarter planet CO and art or



### Netezza Connector – Write Modes

Write Mode	Behavior
Insert	If running in parallel, each processing node has its own external table but they all insert into the same target table.
Update	After all processing nodes are done inserting data into the TWT, the connector executes the update statement from the TWT into the target table once from the conductor process.
Delete	The DELETE statement is generated and it inserts data from the external table into the TWT. After all processing nodes are done inserting into the TWT, the connector executes the delete statement from the TWT into the target table once from the conductor process.
Update then Insert	After all processing nodes are done inserting data into the TWT, the connector first executes the update statement from the TWT into the target table and then executes the insert statement from the TWT into the target table once from the conductor process.
Delete then Insert	After all processing nodes are done inserting data into the TWT, the connector first executes the delete statement from the TWT into the target table and then executes the insert statement from the TWT into the target table once from the conductor process.
Action Column	A char(1) column in the input data determines which SQL statement is executed with the data in that row. The action column can have one of the following values: I for Insert, U for Update, D for Delete, R for Replace (Delete then insert) and M for Merge (update or insert if record does not exist).
User-Defined SQL	Allows for custom SQL to be defined.

for a smarter planet C S C

### Connector Migration Tool – Migrating jobs that use NZ Enterprise

for a smarter planet

- Since the Netezza Enterprise Stage is deprecated, leverage the Connector Migration tool to upgrade jobs to the Netezza Connector
- Windows, Client side tool
  - Located under "Programs", "IBM InfoSphere Information Server" : C:\IBM\InformationServer\Clients\CCMigrationTool\CCMigration.exe
  - GUI or command-line

<ul> <li>Link to Information Center</li> </ul>					Java Transform	er (Parallel)	
•	<ul> <li>Migration tool will analyze all jobs in project</li> <li>Identify jobs to migrate</li> </ul>				Java Client (Par V Netezza Enterp) ODBC (Server) ODBC Correction	allel) ise (Parallel)	
	🖏 Connector Migration Tool File Repository View						OK Cancel
	🔍 Analyze 💣 Migrate 🗄 Preferences						
	ipsvm00079/slbrokaw	Status	Source property	Source value	Target property	Target value	
	🖃 🔄 🔐 JODS	•	server	NZSQL	/Connection/DataSource	NZSQL	
	Netezza Enterprise ET (PyNete		dbname	support	/Connection/Database	support	
	Row Generator (PyRowGenerat		user	team2	/Connection/Username	team2	
	NZEnterprise NZLoad (Parallel Job)		password	{iisenc}V9D/	/Connection/Password	{iisenc}V9D/52dtM	
	Retezza Enterprise nzload (Pxt)		loadmethod	nzload	/Usage/WriteMode	(No equivalent value)	
	Row Generator (PyRowGenerat		operator	nzwrite	(Complex Mapping)	nzwrite	
	NZEnterprise Read (Parallel Job)		mode	append	/Usage/TableAction	0	
	Netezza Enterprise FT (PxNete		table	MyTable	/Usage/TableName	MyTable	
	Peek (PxPeek)		truncate		/Usage/TruncateColumnNames	false	

- Limitations:
  - nzload no equivalent option in Netezza Connector Jobs are migrated to default options for load ET and TWT



### **Balanced Optimization**





- Reduce disk I/O and data movement
- Push processing to database without requiring manual SQL
- Additional details available in <u>developerWorks Article</u>
- Link to Information Center



### Troubleshooting – Debug Environment Variables

Stage	Variable
ODBC Connector	CC_MSG_LEVEL The following list contains the valid values: 1 - Trace 2 - Debug 3 - Informational 4 - Warning 5 - Error 6 - Fatal
Netezza Enterprise Stage	APT_DEBUG_MODULE_NAMES=nzutils odbcconn odbcenv nzwriteop nzsubop nzreadsubop nzupdtsubop nzreadrep nzreadop nzts nzwriterep odbcstmt nzimportustrgf nzimportgf
Netezza Connector Stage	CC_NZ_LOG_LEVEL The following list contains the valid values: 1 - Trace 2 - Debug 3 - Informational 4 - Warning 5 - Error 6 - Fatal This variable is identical to the CC_MSG_LEVEL environment variable, but except that it affects only the Netezza connector messages.

for a smarter planet M



### Troubleshooting – Netezza Log Files

- How to redirect \*.nzoutput, \*.log, \*.nzbad, \*.nzlog files from /tmp
  - Technote <u>144539</u>
  - Netezza Enterprise Stage
    - Set APT\_NETEZZA\_LOAD\_FILES to the desired target directory
  - Neteza Connector Stage
    - Set "Directory for log files" property within stage

Write mode	Insert	
Table name *	MyThirdTable	
Enable case-sensitive identifiers	No	
Truncate column names	No	
SQL		
Table action *	Append	
▼ Session		
<ul> <li>Schema reconciliation</li> </ul>		
Unmatched link column action	Drop	
Type mismatch action	Drop	
Unmatched table column action	Ignore nullable	
Mismatch reporting action	Warning	
<ul> <li>Temporary work table mode</li> </ul>	Automatic	
Table name *		
Truncate table	No	
Drop table	Yes	
Generate statistics	No	
Maximum reject count	1	
Directory for named pipe (Unix only)		
Directory for log files	/home/dsadm/NetezzaLogs	
Other options		

## Troubleshooting – ODBC Driver Manager Tracing

tor a stharter plane!

- Technote <u>1441559</u>
- Compare ODBC trace output with Netezza's Postgres Log
- Netezza Postgres log located:

/nz/kit/log/postgres/pg.log

Windows:

### Use 32 bit ODBC Data Source Administrator

ODBC Data Source Administrator (32-bit)			
User DSN System DSN File DSN Drivers Tracing Connection Pooling About			
When to trace			
Start Tracing Now			
Machine-Wide tracing for all user identities			
Log File Path			
C:\Users\dsadm\ODBC_Trace.log C:\Windows\system32\odbctrac.dll			
Browse Select DLL			
ODBC tracing allows you to create logs of the calls to ODBC drivers for use by support personnel or to aid you in debugging your applications.			
OK Cancel Apply Help			

### Unix:

### Modify the .odbc.ini file

[ODBC] InstallDir=/opt/IBM/InformationServer/Server/branded\_odbc Trace=1 TraceDII=/opt/IBM/InformationServer/Server/branded\_odbc/lib/ VMtrc00.so TraceFile=/tmp/odbctrace.out

Optional: TraceOptions=3 [Display timestamps in trace file] Information Management are for a sharter planet

### Troubleshooting – ODBC Driver Tracing

### Windows:

### Use 32 bit ODBC Data Source Administrator

ODBC Data Source Administrator (32-bit	t) ×
User DSN System DSN File DSN Drivers Tracing Connection Pooling About	]
System Data Sources:	
Name Platform Driver	Add
NZSQL 32-bit NetezzaSQL	Remove
	Configure
IBM Netezza ODBC Driver Setup	×
DSN Options Advanced DSN Options SSL DSN Options Driver Options	
Optimize for ASCII character set	
Prefetch Count: 256	er.
Res	et Defaults
OK Cancel	I Help

### Unix:

Create an odbcinst.ini file in the same location as odbc.ini [.odbc.ini]

Add the following entries:

[ODBC Drivers] NetezzaSQL=Installed

[NetezzaSQL] Driver=/opt/Netezza/lib64/libnzodbc.so Setup=/opt/Netezza/lib64/libnzodbc.so APILevel=1 ConnectFunctions=YYN Description=Netezza ODBC driver DriverODBCVer=03.51 DebugLogging=true LogPath=/tmp/Netezza UnicodeTranslationOption=utf8 CharacterTranslationOption=all PreFetch=256 Socket=16384



**Questions?** 



### Backup Slides – Helpful Links

- Connectivity Guide for Netezza Performance Server
- Whitepaper on Best Practices and Performance Guidelines
- Ensuring transactional consistency with Netezza when using CDC and DataStage
- Additional Netezza Connector Environment Variables