



Utah Oracle Users Group



# *Migration Made Easy*

## *TOS-Talend Open Studio*



# Moving Data to Oracle

Featuring

Sam Weber, OCP

Senior Oracle DBA, Attask Inc.

[sam.weber@attask.com](mailto:sam.weber@attask.com)

*After taking this class, you should  
know how to:*

- Download and install the TOS software
- Understand the common functions of the TOS system
- Create source and destination connections to a variety of dbms and file systems
- Understand and use the tool and functional section panes

*After taking this class, you should  
know how to:*

- Create and run jobs which transfer data
- Map tables and columns from source to dest
- Clean, edit data on the fly
- Replace column data from lookups
- Change data types
- Process records from flat files
- Diagnose and fix common transfer errors

# *What is TOS?*



**Talend Open Studio for Data Integration** is an open source GUI software system which can be downloaded free and used to control the controlled transfer of data from and to a wide variety of DBMS including Oracle.

**TOS** is created and supported by the Talend\* Company. They offer powerful supporting tools to the product which can be purchased and delivered on-line.

**TOS** is solid, well tested and offers a variety of educational materials and programs in how to use the product, many of which are free

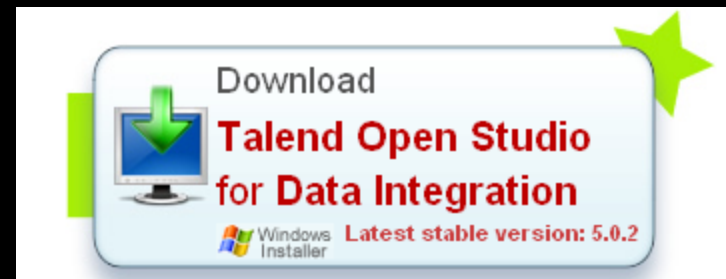
## *Why use TOS?*

- Reliable
- Fast
- Incredible Flexibility
- Broad base of features and functions
- Works from and to every major DBMS
- Main GUI tool is FREE for download
- Can be automated

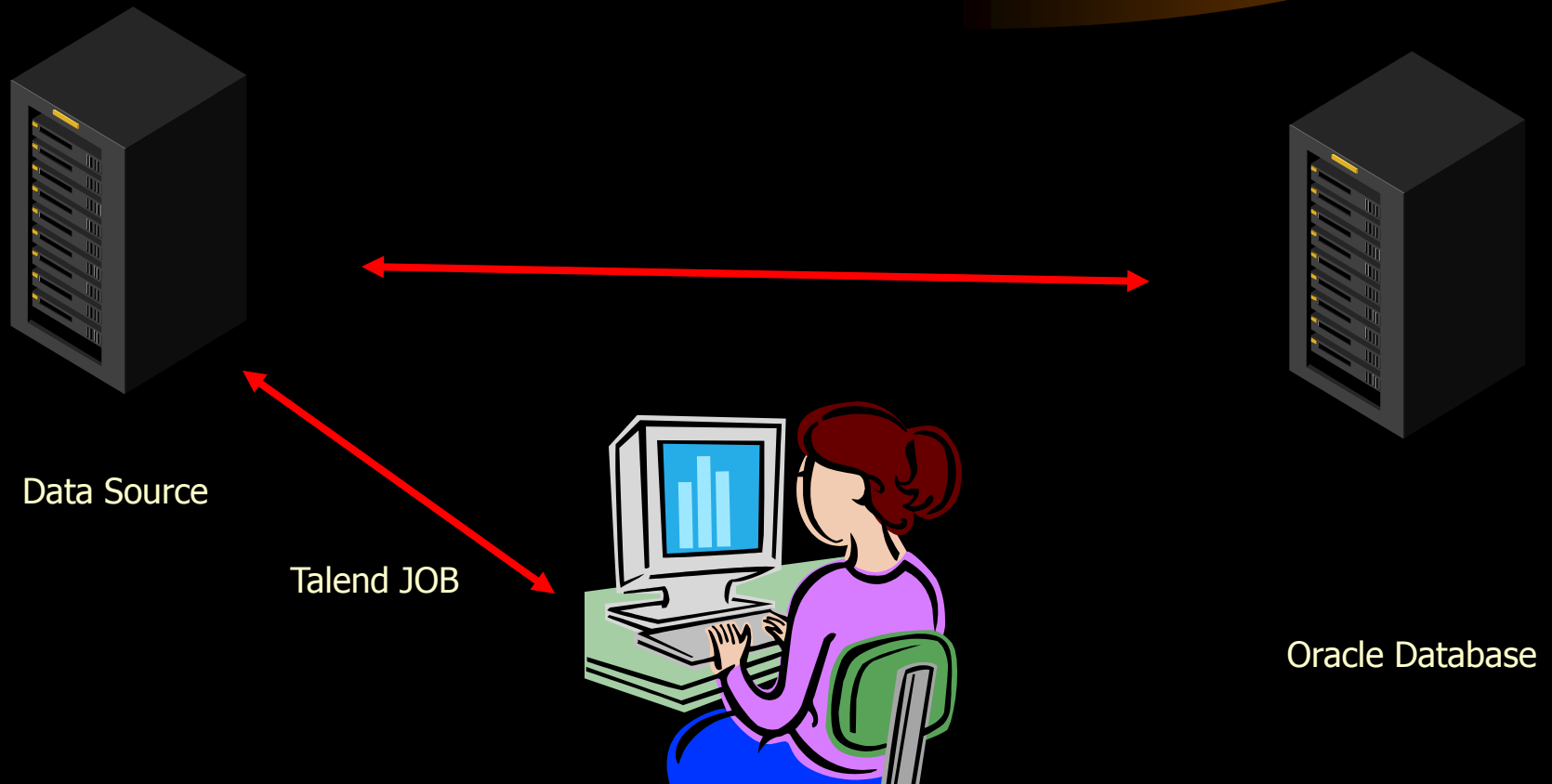
# *How do I get TOS?*

**Just go to their website:**

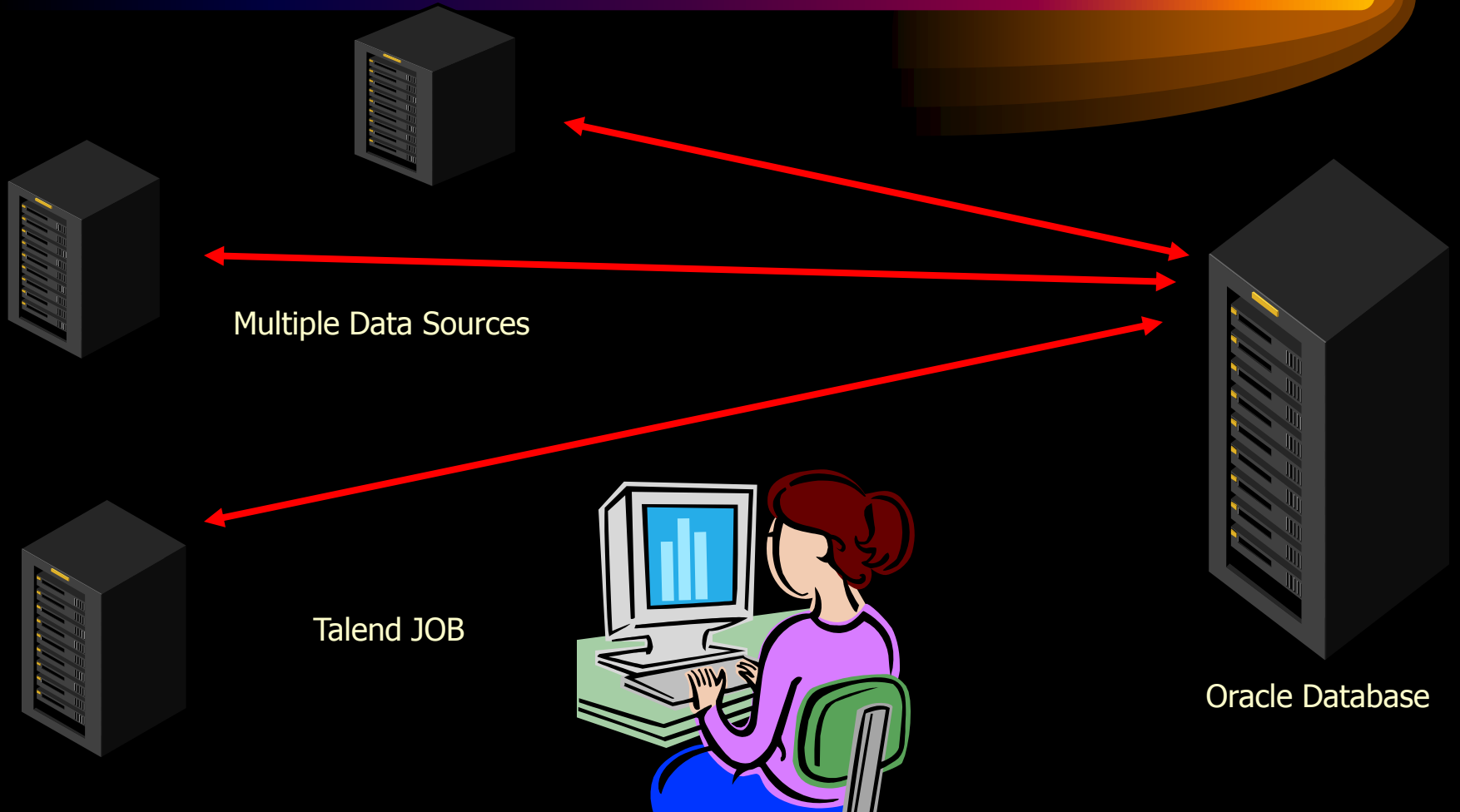
**<http://www.talend.com>**



# *Job from Client transfers data from one DBMS to another*



# *Job from Client transfers data from multiple DBMS to one*





# TOS Sections (Panels)



- Reposit
  - JOB Designs
    - Jobs
  - Metadata
    - DB Connections
    - Files

# TOS Sections (Panels)



- Palette
  - Processing
    - Process Icons
- Run Job
  - Basic
  - Debug
- Component

# *Drag & Drop Component Icons*

Data Source  
Table  
Schema

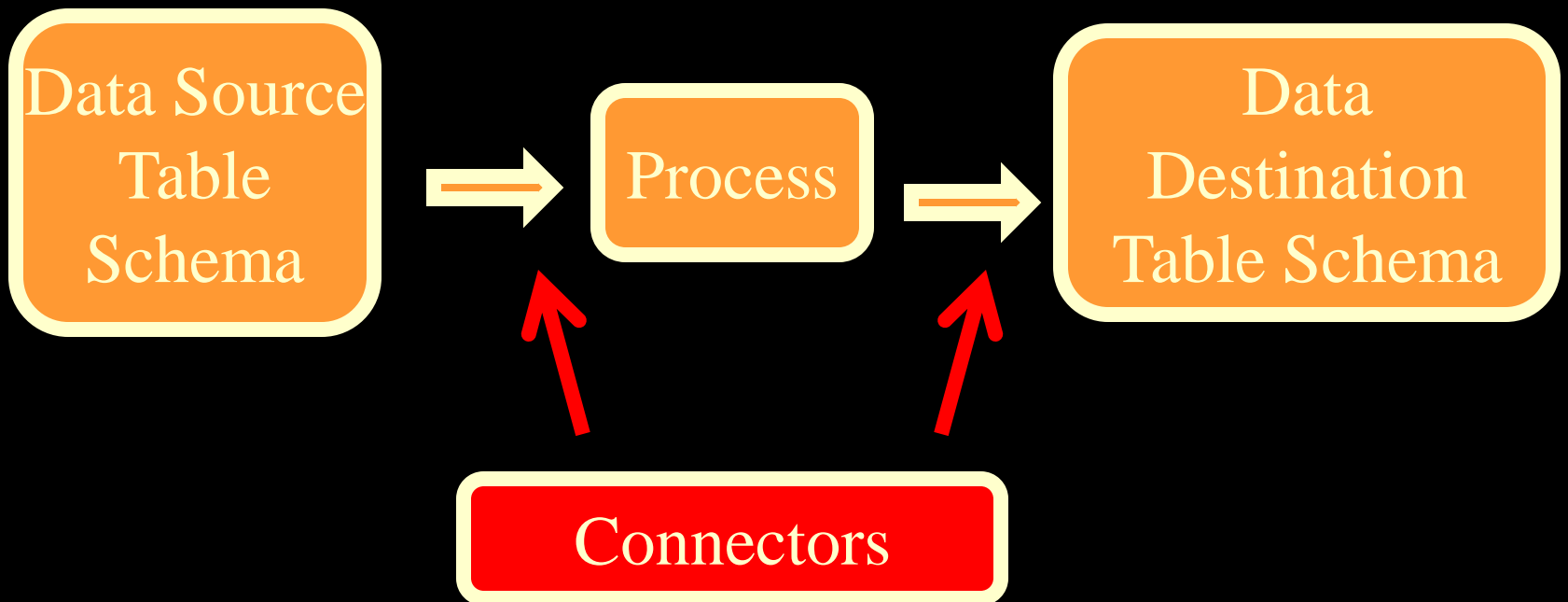
**METADATA**

Data  
Destination  
Table Schema

Process

**PALETTE**

# *Job Build Components*



## *DB Connection*

- Name it as a Source, Dest or both
- 35 different DBMS systems can be connected to
- Default ports come in automatically
- You enter login, password, hostname, db name, schema or additional parameter
- Click button to check connection

## *Retrieve Schema*

- Define search criteria for table names
- Retrieves List of Table, View names
- Mark the ones you want
- Places table/view names in tree view under the connection
- Click icon for table/view – drag and drop it into Create Job pane

## *Flat Files*

- Delimited, Positional, XML, Excel, Regex, Salesforce, LDAP
- Use first row to create logical column names OR define your own
- Format files for Windows, Linux or Mac
- Viewer shows subset of data
- Skip specific rows or limit total number

## *The Job Pane*



- Free Space to drag icons to
- Easily resized to hold one or hundreds of individual jobs
- Magnify or Reduce
- Dragging a data connection in requires secondary definition



# *Palette – Processing icons*



- tMap
  - Column to column mappings
  - Auto map for columns with same name
  - Change Data Type
- tFilterColumns
  - Remove unwanted columns
- tFilterRow
  - Remove rows based on column values

## *Palette – Processing icons*



- **tReplace**
  - Substitute column values as data travels across
- **tJoin**
  - Substitute a lookup value from different source
- **tSortRow**
  - Sort for column(s) value ASC or DESC

## *Palette – Processing icons*

- `tAggregateRow`
  - Perform count, sum, min, max on column values to return single aggregate
- `tConvertType`
  - Change datatype via formula or function
- `tSampleRow`
  - limits the result set to a specific number of rows

## *The Component Pane*

- Appears when input/output icon is highlighted
- Displays connection name, host, db, schema, port and table name
- Action on table allows the dest table to be created, dropped, cleared of data or truncated before data inserts
- Validates the connection

## *Run Job pane*



- Single click compiles and runs job to move the data
- Job displays the row transfer speed and total number of rows created
- Fatal compile or execution errors display just above the icon
- Can run in normal or debug mode

## *Job execution result pane*



- Displays log of execution results
- Logs the individual row failure errors
- Brings back description of error
- Shows successful completion of all jobs

## *Contact Info*



- [sam.weber@attask.com](mailto:sam.weber@attask.com)
- Contact me if you have follow-up questions
- Check Out [www.attask.com](http://www.attask.com) on the internet, it's a great Project and Portfolio management software for project managers, developers and executives



*Questions?*



Utah Oracle Users Group

