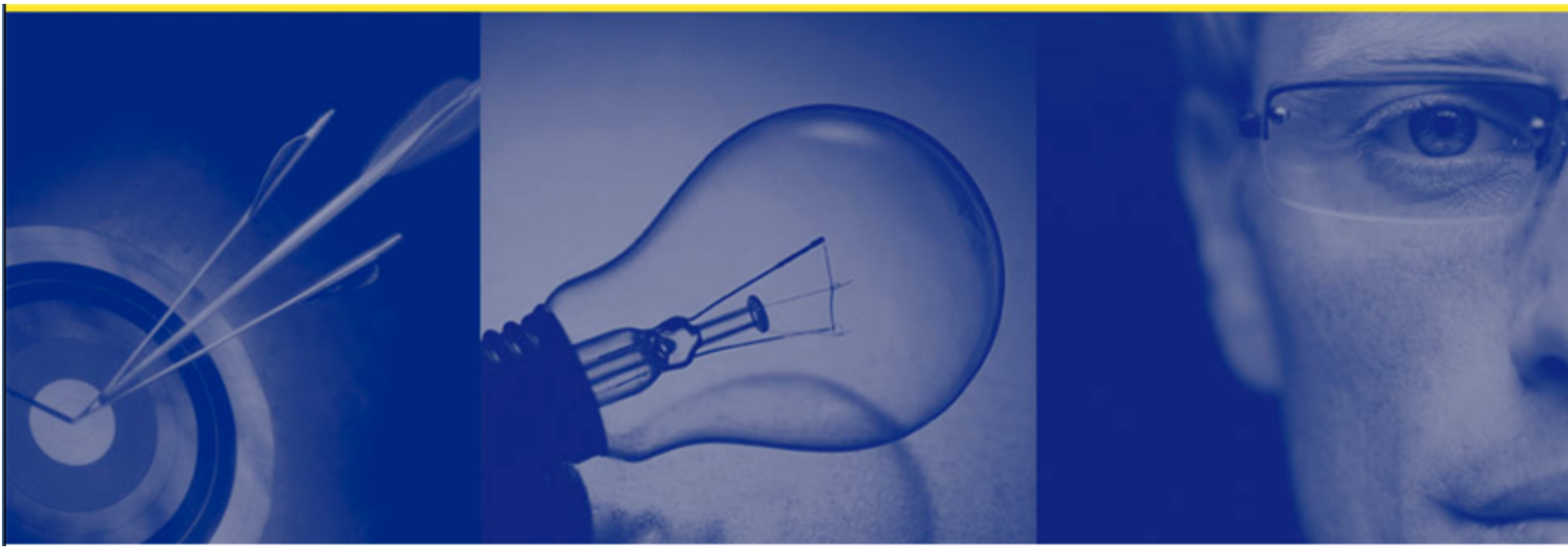
QUEST SOFTWARE

# Toad® for Cloud Databases 1.1

## Getting Started Guide



Quest Software World Headquarters
LEGAL Dept
5 Polaris Way
Aliso Viejo, CA 92656

www.quest.com
email: legal@quest.com

Toad for Cloud Databases 1.1
Getting Started Guide
October 2010

# Table of Contents

# 1

# About Toad for Cloud Databases

## What does Toad for Cloud Databases do?

What is a Cloud database? Cloud databases:

- Process huge volumes of data. The volume of data produced by a web site like Google, Twitter and Facebook for example is too enormous to be processed by a conventional relational database. At the time of writing this guide the daily message volume of Twitter was around 7 terabytes per day.
- Host data in the Internet cloud. Cloud databases are accessed via the Internet.
- Are also called NoSQL databases because their interface is not native to SQL. To query, access and modify the data requires code that can be tedious, expensive and time consuming even for software developers to write.

Toad for Cloud Databases enables users to use SQL to:

- Work with data in a Cloud database.
- Join a Cloud database to a relational database on the personal computer.
- Copy tables from one Cloud database to another or between a Cloud database and a relational database on the personal computer.

SQL is an industry standard commonly used by data analysts, data administrators and business intelligence (BI) analysts. Quest Software's Toad product is a powerful data query and administration tool used extensively within the Oracle community and in recent years diversified to encompass SQL Server, MySQL, DB2 and other relational databases.

## How does Toad for Cloud Databases work?

There are two key components to Toad for Cloud Databases. The first is the Toad client. It is installed on a Microsoft Windows computer. It is used to access the Cloud database and write and execute SQL statements. Features common to Toad clients such as the SQL editor, visual query builder, project management and reporting facilities are available.

The second component is the Data Hub. It is a data server installed with the Toad client. It can be run from the local computer or deployed in the Amazon cloud. The Data Hub translates SQL statements submitted through the Toad client into a language understood by the Cloud database, and returns results in the familiar tabular row and column format.

## Which Cloud databases are supported?

The following Cloud database technologies are supported:

| Technology | Service Provider | For further information |
|---|---|---|
| Azure Table Services | Microsoft | Introduction to Windows Azure<br>Table Services Concepts |
| Cassandra | Apache Open-Source | Apache Cassandra home page |
| SimpleDB | Amazon | Home page<br>Getting Started Guide |
| Hadoop HBase | Open-Source: based on Google technologies | Home page |
| ODBC-compliant relational databases | | Though not a Cloud database, support for ODBC-compliant relational databases enables Toad to join data from a local relational database to a remote Cloud database in a single SQL query. |

# 2

# Install Toad for Cloud Databases

## System Requirements

Before installing Toad, ensure that your client system meets the following minimum hardware and software requirements:

| | |
|---|---|
| **Platform** | 1 GHz |
| **Memory** | 512 MB of RAM minimum, 1 GB recommended |
| **Hard Disk Space** | 550 MB |
| **Operating System** | Windows XP (32-bit)<br>Windows Vista (32-bit)<br>Windows 7 (32-bit and 64-bit) |
| **.NET Framework** | Microsoft .NET Framework 3.5 (Service Pack 1)<br>**Note:** The following conditions apply:<br>• The .NET Framework Client Profile is not supported.<br>• To run Toad, the .NET security policy must be set to unrestricted. See "User Requirements and Restrictions for Running Toad" (page 9) for more information. |
| **Additional Requirements** | **Web Browser**<br>Microsoft Internet Explorer 6.0 or later |

## User Requirements

### User Requirements to Install Toad

Only Windows Administrators and Power Users can install Toad.

If an Administrator installs Toad, the Administrator can share the Toad shortcut icon on the desktop and menu items with any user--a Power User, Restricted User, Vista standard user--on the computer.

If a Power User installs Toad, a Restricted User cannot do the following:

- Associate file extensions.
- View the Toad shortcut icon on the desktop or the Toad menu options in the Windows Start menu. As a workaround, Restricted Users can create a shortcut to Toad from the install directory and place the shortcut on their desktop.

### User Requirements and Restrictions for Running Toad

The following describes the requirements and restrictions for users who launch and run Toad:

- To run Toad, the .NET security policy must be set to *Unrestricted*. To modify the security policy:
    1. Select **Control Panel** from the Windows **Start** menu.
    2. Double-click **Administrative Tools**.
    3. Double-click **Microsoft .NET Framework *version* Configuration**.
    4. Select **My Computer**, and then select the **Runtime Security Policy** node.
    5. Select **Evaluate Assembly** from the list of tasks in the right panel, and locate and select Toad.exe in the **File** field.
    6. Click **Next** and set the permission to *Unrestricted*.
- To create or modify the QuestSoftware.ToadSecurity table to enable Toad Security, the user who maintains this table must have privileges to create, insert, update, and delete it. All other users need only the ability to read the QuestSoftware.ToadSecurity table to use Toad Security features.

## Install Toad

***To install Toad***

1. Run the *ToadInstaller.exe*.
2. Complete the wizard.

## Install ODBC Drivers

On selecting the desktop deployment model install suitable ODBC drivers on the desktop in order to connect to ODBC data sources.

Recommended drivers for the major RDBMSs are:

- Oracle: Oracle 11g instant client Basic + ODBC Supplement
- SQL Server, SQL Azure: SQL Server Native Client 10.0
- MySQL: MySQL Connector/ODBC 5.1

# Uninstall Toad

If you uninstall Toad, it may not uninstall completely. When you use Toad, you create new files such as SQL scripts, data output, or temp files. The following types of files are not removed when you uninstall:

- Any new files you create
- Any copies of files you make for backup purposes
- The desktop DataHub deployment, in particular the *DataHub* folder and its *conf* and *data* sub folders.
- For Windows Vista and Windows 7 any data related to Toad in the VirtualStore in %LOCALAPPDATA%

### *To uninstall Toad*

1. Run the *ToadInstaller.exe* installation program from the CD or downloaded file.
2. Select **Uninstall Toad only**.
3. Complete the wizard.

# 3

# From Remote Data Source to SQL Tables

## Deploy Data Hubs

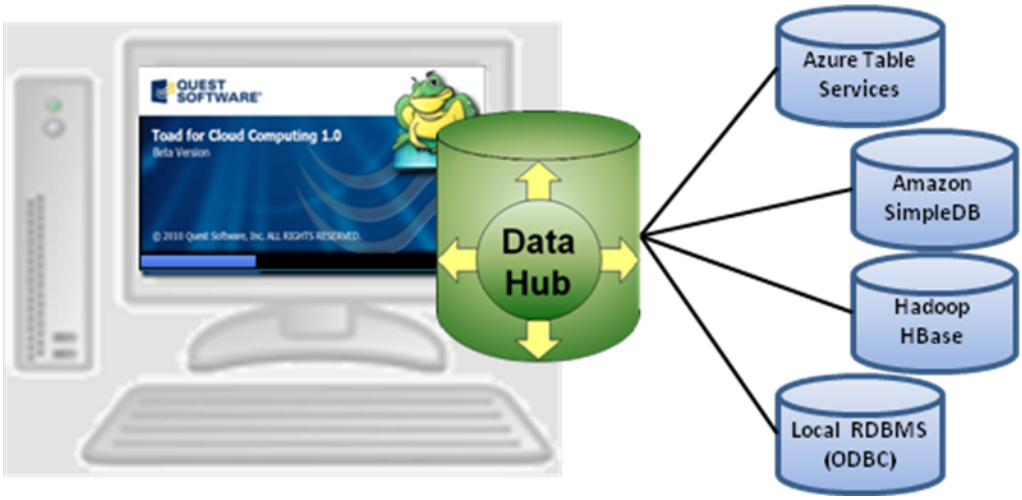
Once a Data Hub is deployed then a data source can be mapped. Use the following table to determine which of the two Data Hub deployment models is appropriate to your requirements.

| Deployment Model | Where is the Data Hub? | Who can access the Data Hub? | What does this mean for data access? |
|---|---|---|---|
| Desktop | The Data Hub is local to the Toad Client. It is installed automatically with the Toad client. | The Data Hub is private. The Toad client that deployed the Data Hub has exclusive access to it. | Queries to the data source are exclusive to one Toad client. |
| Cloud | The Data Hub is created in an Amazon EC2 instance in the cloud. | The Data Hub may be shared by multiple Toad clients. | Queries to the data source may be setup by one client, then shared by many Toad clients. |

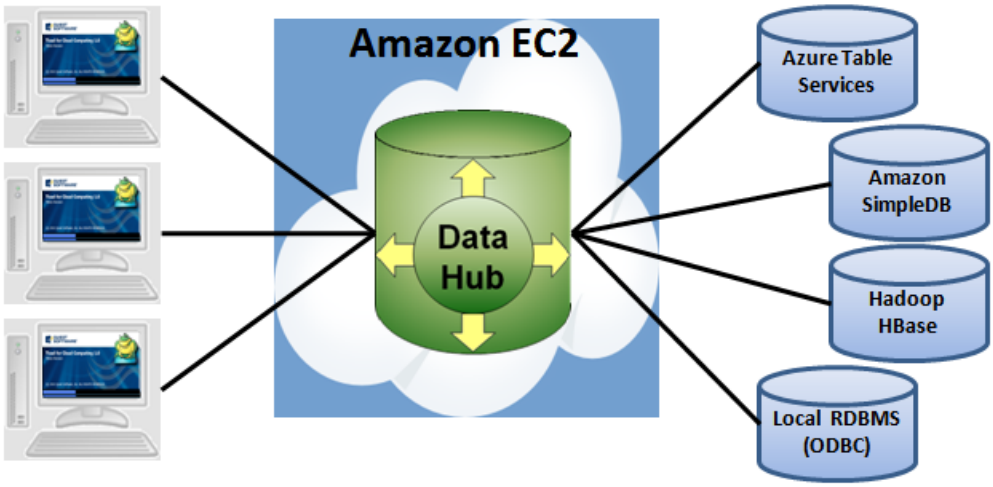
**Notes:**

- The Toad client can switch between the desktop deployment and multiple Cloud deployments as required.
- To deploy an additional Data Hub click **Tools** | **Data Hub Deployment Wizard**.
- Toad incorporates an Amazon EC2 console to support use of the cloud deployment model. See "Cloud Deployment and the Amazon EC2 Console " (page 20) for more information.

## The Desktop Deployment Model




## The Cloud Deployment Model




# Map Data Sources

Toad prompts the user to map the data source following deployment of the Data Hub and while the Data Hub has no mappings.

Configure as prompted:

| Option | Description |
|---|---|
| Data source name | Enter a free-form name to the data source. This name is used for display purposes in Toad. |
| Data source type | Select a data source type from the list. |
| Parameters of the data source | The parameters vary with the data source type.<br>• Azure — Enter the username and password to connect to the Azure source.<br>• Cassandra — Enter the host, port, keyspace, username and password to connect to the Cassandra source.<br>• HBASE — Enter the URI of the HBASE data source.<br>• ODBC — Select the database type. Enter parameters as appropriate to that database type. Additional information is available for:<br>  • SQL Azure — See "Map to a SQL Azure Data Source" (page 13) for more information.<br>  • SQL Server — See "Map to a SQL Server Data Source" (page 14) for more information.<br>• SimpleDB — Enter the username and secret key to connect to the SimpleDB source. |

**Note:** To map an additional data source, click **Data Sources** | .

## Map to a SQL Azure Data Source

To map to a SQL Azure data source, select ODBC as the data source type and enter the parameters as follows.

| Parameter | Value |
|---|---|
| Data source type | ODBC |
| Database type | SQL Azure |
| Driver name | “SQL Server native Client 10.0” (or higher) |
| Server name | The name of the server.<br>**Note:** Do not prefix the server name with *tcp:* |

| Parameter | Value |
|---|---|
| Database name | The name of the database. |
| User ID | The User ID to connect to the database.<br>**Note:** Do not include the server name with the User ID. If the connection string is of the form user_id@server, then in the User ID field just type the user_id. |
| Password | The Password to connect to the database. |

## Map to a SQL Server Data Source

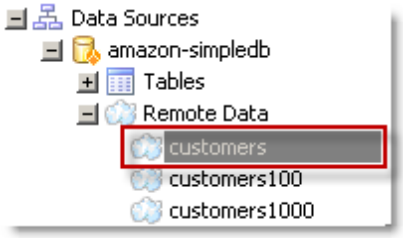
To map to a SQL Server data source, select ODBC as the data source type and enter the parameters as follows.

| Parameter | Value |
|---|---|
| Data source type | ODBC |
| Database type | SQL Server |
| Driver name | “SQL Server native Client 10.0” (or higher) |
| Instance name | The name of the database instance. |
| Database name | The name of the database. |
| Use Windows integrated security | Select to use Windows integrated security.<br>**Note:** When not selected, enter a value for the User ID and Password. |
| User ID | The User ID to connect to the database.<br>**Note:** A value is required when Windows integrated security is not selected. |
| Password | The Password to connect to the database.<br>**Note:** A value is required when Windows integrated security is not selected. |

## Map Data Source Objects To SQL Tables

Within the data source are organizational units similar to an SQL table. For Azure, HBASE and ODBC they are called tables. For SimpleDB they are called domains. To work with a remote organizational unit in Toad it must be mapped to an SQL table.

1. Click **Data Sources** | *the data source* | **Remote Data** | *the data object* in the **Object Explorer**.



2. Click **Map Remote Data**.
3. Configure as required:

| Option | Description |
|---|---|
| Data source | The data source type. This was defined when the data source was mapped. See "Map Data Sources" (page 12) for more information. |
| Remote data object (That this table will map to) | By default this is the selected data object. Any data object in the data source can be selected.<br>**Note:** To filter the list of data objects click . |
| Table name | The name of the table as it appears in Toad. By default the name of the remote data object is also the table name. |
| Table fields | By default all fields in the data object are mapped to the table. Field names and data types are as per the data object.<br>The dialog shows:<br>• Remote Field — The name of the field in the remote data object.<br>• Local Name — The name of the field in the local table.<br>• Datatype — The data type of the field in the local table.<br>• Not Null — Selected when Null values are not allowed in this field of the local table.<br>Actions can be taken to choose which fields to map, rename a field and change a field's data type and required status.<br>**Caution:** Issues can arise when taking action on a field with an *identifier* data type. See "Taking Actions on Fields with identifier Data Type" (page 18) for more information.<br>***To not map a field:***<br>a. Select the field to remove.<br>b. Click **Remove**.<br>**Note:** After fields have been removed, you can restore them to the list with the **Add** function.<br>***To rename a field:***<br>a. Click the name in the **Local Name** list.<br>b. Rename.<br>***To change a field's data type***<br>Select an alternative. |

| Option | Description |
|---|---|
| | **Caution:** The Map Remote Data function fails if the data does not match the data type. For example, setting the field's data type to *integer* causes Map Remote Data to fail if the field has non-numeric data.<br><br>***To change a field's required status***<br>Click **Not Null**. When Not Null is selected a data value is required in each record of the field.<br><br>**Caution:** The Map Remote Data function fails if Not Null is selected for a field that has null data. |

## Considerations

### Taking Actions on Fields with *identifier* Data Type

Columns with name *id* and data type *identifier* are generally part of the table's metadata and should not be removed.

| Data source | Organizational unit | Number of columns with data type identifier |
|---|---|---|
| Azure | Table | 2 |
| Cassandra | Column Family | 1 |
| HBASE | Table | 1 |
| SimpleDB | Domain | 1 |
| ODBC | Table | 1 or more |

### When the Data Source is an Oracle Database

- Ensure that for the Toad table the data object name and all field names are in UPPERCASE. The Data Hub receives column metadata from Oracle in uppercase, regardless of the letter case the metadata was created in.
- Toad does not recognize a 64-bit version Oracle ODBC on a 64-bit Operating System. Use a 32-bit version of the Oracle client (Oracle server, client or instant client [Basic and ODBC] ).

## When the Data Source is a Cloud Database

Internet HTTP access is required to access a Cloud database. When there is no Internet HTTP access the Data Hub may experience connection time outs on attempts to map a data source to a Cloud database.

If Internet HTTP access is through a HTTP proxy server then setup is required before the data source is mapped.

1. Close Toad
2. Set the environment variable that will route all HTTP traffic through the proxy server port.
    a. The name of the environment variable is *http_proxy*.
    b. The value of the environment variable is of the form http://*proxyAddress*:*proxyPort*

       For authenticated proxies the value of the environment variable is of the form http://*username*:*password*@*proxyAddress*:*proxyPort*

        a. The *username* and *password* are required for authenticated proxies. They are the username and password to the proxy server within the user's organization.
        b. The *proxyAddress* is the DNS or IP address of the proxy server within the user's organization.
        c. The *proxyPort* is the port on which the proxy server operates. The standard proxy port number is 8080. Sometimes port 80 is used. Less commonly an entirely different port may be used.
3. Re-start Toad.
4. Map the data source.

**Note:** If you see an error like *Microsoft Azure requires client applications to use HTTP 1.1* then it is likely your proxy server has converted the Data Hub to use HTTP version 1.0. By default the Data Hub uses HTTP version 1.1.

## When the Data Source is HBase or Cassandra

The mapping of remote data objects to tables is more complicated for HBase and Cassandra. These two NoSQL databases differ from the others in that they both utilize a column family construct for data organization. In essence, a column family is a grouping element that sits between the level of the table and the column, and groups one or more related columns together.

A full explanation of column families in both these NoSQL databases can be found on the Toad for Cloud Database WIKI page in the two articles Column Families 101 for HBase and Cassandra Column Families.

The pertinent consideration here is the effect that column families have on the mapping of NoSQL data. Toad for Cloud Databases maps dynamic column families as sub-tables. We use the term dynamic to denote that either the family names or column names are actually data rather than static identifiers. In these cases, the columns in the column family are abstracted out to a child table that is subordinate to the main table that is being mapped. The main table and its sub-tables can then be relationally joined using the identifier fields from the main table.

To alleviate some of the confusion that is bound to arise from this sometimes complicated mapping process, Toad for Cloud Database has introduced a Preview button in the mapping dialog. Clicking the Preview button will show the user what the current table and sub-table mappings will look like using real data from the remote data source. This makes it easier for the user to see where they may need to change the mapping information to create the desired relational tables.

## Cloud Deployment and the Amazon EC2 Console

The Toad Amazon EC2 console provides a way to manage a cloud deployment of the Data Hub internal to Toad. See "Deploy Data Hubs" (page 11) for more information.

Advantages of the Amazon EC2 console:

- Toad's EC2 console provides a single integrated view of EC2 instances from which all common administrative tasks can be performed. Other EC2 management tools are component-oriented. They do not focus on user workflows. They may have facilities for managing components such as instances, AMIs, security groups, storage volumes, etc, but their facilities are segregated into different tabs in a web browser. Toad adds facilities such as a favorites list for Amazon machine images (AMI) so there is no need to wait for thousands of AMIs to be retrieved to launch an instance of a frequently used image.
- A future goal of the Amazon EC2 console is to support workflows that include multiple steps. Such workflows will enable potentially complex deployments to be executed with a single step. An example deployment scenario may be to launch a Data Hub instance then install extra ODBC drivers. A more complex scenario may involve deploying an Oracle database instance in the cloud with specific network configurations, service names, users and tables.

To open the Amazon EC2 console

- Click  on the Toad tool bar.

Use the Amazon EC2 console according to the following table. If prompted to link to an Amazon EC2 account See "Link to an Amazon EC2 Account" (page 23) for more information.

| Window Section | Description |
|---|---|
| Instance List | All virtual machine instances running under the selected Amazon EC2 account are listed. Select an instance. The *Instance Details* and *Volumes* sections of this panel are updated accordingly. |

| Window Section | Description |
|---|---|
| | **Tips:**<br>• Click **Refresh** to redraw the list of instances.<br>• Click **Get windows password** to obtain the Administrator password for the selected instance. This is applicable to instances on a Windows platform. The user requires this information to connect to the Windows instance using a remote desktop application.<br>• Click **New instance** to create a new instance from a static Amazon machine image.<br>  ◦ At the New Instance window, select an AMI to use as a template for constructing the instance. The details of the selected AMI are displayed in the information panel.<br>  ◦ It is recommended to use the Favorites tab of the New Instance window. From the Amazon Images tab the entire list of Amazon AMI's are available. Select an AMI of interest and click **Add to favorites**. Repeat for all Amazon AMI's of interest. AMI's added to favorites can be selected and browsed from the Favorites tab.<br>  ◦ To search for an AMI on the Amazon Images tab, type some text to filter on and click **Search**. For example, type hbase to search for all AMI's containing the keyword hbase.<br>• Click **Shut down instance** to terminate the selected instance. |
| Instance Details | A display of the properties of the machine and the Amazon image from which the selected instance was launched. Use this information to create another instance of the same type, or verify the exact version and type of the running instance from the image's manifest. The display includes the following:<br>• Instance ID -Uniquely identifies the instance among all EC2 instances.<br>• Public IP - Use this IP address to connect to this instance from outside the Amazon cloud<br>• Image ID - Uniquely identifies the image among all AMIs.<br>**Tips:**<br>• To update the Instance Details panel, re-select the instance.<br>• Enter a name in the **Display name** field and click **Update Display Name** to help you locate this instance again in the Instance List. |

| Window Section | Description |
|---|---|
| Volumes | A display of the storage devices attached to the selected instance.<br>The display includes the following:<br>• Volume ID – Uniquely identifies the volume among all EBS volumes<br>• Attach Time – The time at which the volume was attached to the selected instance<br>• Device – The disk to which the volume was mapped (for selected instances on the Linux platform)<br>***To attach an EBS volume to the selected instance***<br>An EBS volume - Amazon Elastic Block Storage - is a mechanism for extending storage beyond that available through the initial instance.<br>1. Click **Attach**<br>2. Either:<br>a. Select a volume from those available (and for a Linux platform type the Device)<br>b. Click **Create** to create a volume. Type the size of the volume required. The maximum size for a new volume is 1,000 gigabytes.<br>**Note:** An EBS volume persists until it is deleted by the user, regardless of whether or not it is attached to a running instance.<br>**Tip:** To detach a volume, select the volume and click **Detach**. It may take several minutes to detach the volume. After confirming the operation, the volume will be detached in the background. It will disappear from the list of attached volumes once Amazon completes the operation. |
| Security Groups | A display of the Amazon security groups attached to the selected instance.<br>These security groups determine what network access is allowed to the instance. The allowed access is an accumulation of the rights of each separate security group assigned to the instance.<br>1. Select a security group from the drop down list. All security groups associated with the selected instance are in the drop down list.<br>2. The open ports for the selected security group are displayed in the IP and port range grid.<br>**Note:** Update security information can be updated using the Amazon Web Service portal or the Mozilla ElasticFox plug-in. The Security Groups pane is for display purposes only. |

### Link to an Amazon EC2 Account

Toad prompts the user to link to an Amazon EC2 account on opening the EC2 console if no account is linked.

To link to an additional Amazon EC2 account

1. Open the Amazon EC2 console.
2. Click **Link account**

| Field | Description |
|---|---|
| Account Name | A free-form name to the Amazon EC2 account. This name is used for display purposes in Toad. |
| Access Key | The access key as supplied by Amazon on creating the EC2 account. |
| Secret Key | The secret key as supplied by Amazon on creating the EC2 account. |

## Considerations

### Your Computer Firewall

For Toad to connect to the Amazon EC2 it must be able to make an outgoing connection to port 3306 on the public DNS or IP address of the Data Hub's EC2 instance.

This port may not be open in the internet firewall at the user's site. A request to open this port to the range of Amazon EC2 IP addresses may be problematic in an extremely secure environment. The fact that it is an outgoing connection to a well known set of IP addresses should be acceptable in most environments – especially those that are wanting to explore cloud computing.

The list of Amazon EC2 IP address ranges in the United States as at November 2009 is as follows

| Network Mask | Actual IP Range |
|---|---|
| 216.182.224.0/20 | 216.182.224.0 - 216.182.239.255 |
| 72.44.32.0/19 | 72.44.32.0 - 72.44.63.255 |
| 67.202.0.0/18 | 67.202.0.0 - 67.202.63.255 |
| 75.101.128.0/17 | 75.101.128.0 - 75.101.255.255 |
| 174.129.0.0/16 | 174.129.0.0 - 174.129.255.255 |
| 204.236.224.0/19 | 204.236.224.0 - 204.236.255.255 |

**Note:** This list was provided by an Amazon Web Services (AWS) developer in an AWS development forum. Unfortunately, there is no definitive Amazon list. Any list that is published is always heavily qualified with *may change without notice*.

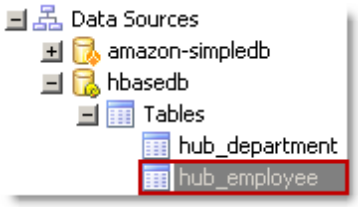
# 4

# Work with SQL Tables in Toad

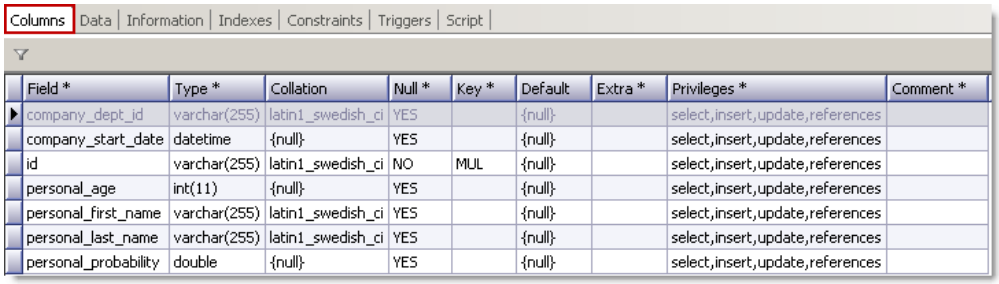
## View Object Details

Once data source units are mapped to SQL tables, the tables are listed in **Object Explorer**. Treat them like relational tables with respect to normal Toad functions.

1. Double click on the table in Object Explorer.

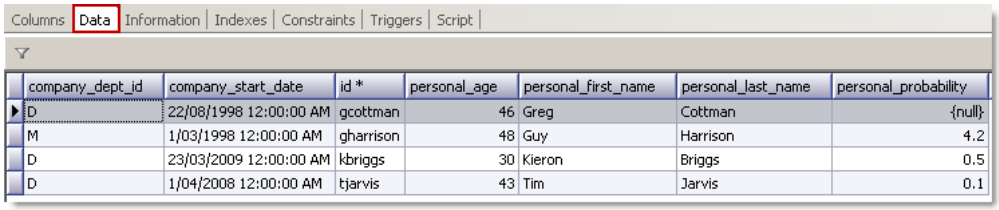



   By default the table opens at **Columns** for a view of the SQL structure.




| Field * | Type * | Collation | Null * | Key * | Default | Extra * | Privileges * | Comment * |
|---|---|---|---|---|---|---|---|---|
| company_dept_id | varchar(255) | latin1_swedish_ci | YES | | {null} | | select,insert,update,references | |
| company_start_date | datetime | {null} | YES | | {null} | | select,insert,update,references | |
| id | varchar(255) | latin1_swedish_ci | NO | MUL | {null} | | select,insert,update,references | |
| personal_age | int(11) | {null} | YES | | {null} | | select,insert,update,references | |
| personal_first_name | varchar(255) | latin1_swedish_ci | YES | | {null} | | select,insert,update,references | |
| personal_last_name | varchar(255) | latin1_swedish_ci | YES | | {null} | | select,insert,update,references | |
| personal_probability | double | {null} | YES | | {null} | | select,insert,update,references | |

2. Click **Data** to view the record values.



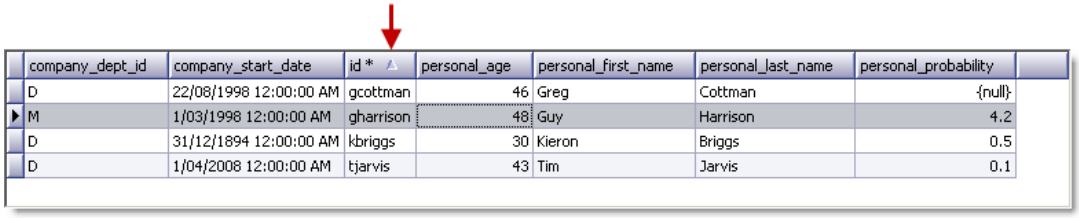

| company_dept_id | company_start_date | id * | personal_age | personal_first_name | personal_last_name | personal_probability |
|---|---|---|---|---|---|---|
| D | 22/08/1998 12:00:00 AM | gcottman | 46 | Greg | Cottman | {null} |
| M | 1/03/1998 12:00:00 AM | gharrison | 48 | Guy | Harrison | 4.2 |
| D | 23/03/2009 12:00:00 AM | kbriggs | 30 | Kieron | Briggs | 0.5 |
| D | 1/04/2008 12:00:00 AM | tjarvis | 43 | Tim | Jarvis | 0.1 |

**Note:** This data viewer is limited to showing the first 1,000 rows of the table.

# Filter the Records on View

***Sort records by ...***

1. Click the title bar of the column to sort by.
2. The column is sorted ascending .
3. Click the title bar again to sort the column descending .



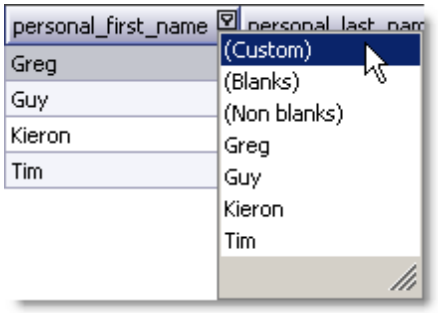
| company_dept_id | company_start_date | id * | personal_age | personal_first_name | personal_last_name | personal_probability |
|---|---|---|---|---|---|---|
| D | 22/08/1998 12:00:00 AM | gcottman | 46 | Greg | Cottman | {null} |
| M | 1/03/1998 12:00:00 AM | gharrison | 48 | Guy | Harrison | 4.2 |
| D | 31/12/1894 12:00:00 AM | kbriggs | 30 | Kieron | Briggs | 0.5 |
| D | 1/04/2008 12:00:00 AM | tjarvis | 43 | Tim | Jarvis | 0.1 |

**Note:** For more complex sorts click |**Order By**.

**Tip:** When the column's datatype is *varchar* or *identifier* then by default values are sorted by lexical order. If the values are numeric then to sort by numeric order use the SQL editor to ORDER BY (column_name AS DECIMAL).

***Show records where ...***

1. Place the cursor over the column title bar. Click .
2. Select from the options to limit the view to records that equal a value, empty records, records that are not empty or create or select a customised where clause.




**Note:** For more complex filters click |**Where Clause**.

***Don't show columns ...***

1. Click .
2. Click **Included Columns**.
3. De-select columns to hide them from view.

# Insert, Update and Delete records

Toad is a powerful data query and administration tool used extensively within the Oracle community and in recent years diversified to encompass SQL Server, MySQL, DB2 and other relational databases.

Use the Toad Query Builder and SQL Editor to write SQL to insert, delete and update records.

## The Query Builder

The Query Builder enables you to create a query without writing or editing SQL statements. Even if you are familiar with SQL, the graphical interface makes it easier to create relationships and visualise the query.

To build a query

1. Click  (CTRL-SHIFT-Q).
2. Drag tables and views from the Object Explorer to the Diagram pane.
3. Select the type of statement you want to create.

## The SQL Editor

The editor combines a powerful SQL and procedure editor into a single interface.

To edit SQL

- Click  on the toolbar (ALT+T+E).

# Copy Tables

A mapped table can be copied to any mapped data source in the Data Hub.

Examples are:

- Copy a local RDBMS table to a Cloud database in development of a Cloud application.
- Copy a mapped table from a Cloud database to a local data warehouse.
- Copy a mapped table from one Cloud database to another.
- Copy tables between local RDBMS.

**Caution:** This release of Toad for Cloud Databases copies data between databases in a single atomic operation. This approach is most efficient for small or medium sized tables, but is not always practical for large or massive data sets. Specifically, there is no way for Toad to provide feedback during the copy operation and – in the case of copies to relational databases – there may be significant demands on rollback/undo segments or transaction logs

during the operation. A future release of the product will provide a copy facility which supports incremental, restartable data transfers with feedback during long operations.

## How to Copy Table

1. Select the table to copy.
2. Click **Copy Table**.
3. Select the **Target data source**. The table is copied to the target data source. Choose from any mapped data source in the Data Hub.
4. Configure as follows:

| Option | Description |
|---|---|
| Target remote data object | The table is copied to this object. By default the object has the same name as the table to be copied. |
| Target column family | Data source technologies that support column families usually require this.<br>For example: When copying to HBase a column family name is required. |
| Target table | The copied table in Toad has this name. By default it is the same name as the table being copied. |
| Source Column / Target Column / Target Remote Field/Data Type / Not Null | By default all columns are copied, their names and data type are unchanged. There may be reason to configure the following fields. See "Considerations on Copy Table" (page 31) for more information.<br>• Source Column — The name of the column or field in the table being copied.<br>• Target Column — The name given to this column (or field) in the copied table in Toad.<br>• Target Remote Field — The name given to this column (or field) in the copied table in the remote data source.<br>• Datatype — The data type of this column (or field) in the copied table.<br>• Not Null — Selected when Null values are not allowed in this column (or field) in the copied table.<br>Actions can be taken to choose which fields to copy, rename a field and change a field's data type and required status.<br>**Caution:** Ensure you are familiar with the target data source requirements when taking action on a columns with name *id* and data type *identifier*.<br>***To prevent one or more columns from being copied:*** |

| Option | Description |
| --- | --- |
| | a. Select the column to remove.<br>b. Click **Remove**.<br>**Note:** After columns have been removed, you can restore them to the list with the **Add** function.<br>***To rename a column:***<br>Columns may need to be renamed, particularly when copying between technologies with different naming conventions.<br>a. Click on the name in Target Column.<br>b. Edit the old name and create a new name.<br>***To change a column's datatype***<br>a. Select the column.<br>b. Select a new datatype.<br>**Caution:** The copy operation fails if the data does not match the new data type. For example, changing the data type to *integer* causes copy to fail if the column has non-numeric data.<br>***To change a column's required status***<br>a. Select the column.<br>b. Select **Not Null**.<br>**Caution:** The copy operation fails if Not Null is selected for a column that has null data. |

| Option | Description |
|---|---|
| Filter copied rows | By default all records are copied, up to the number specified in the *Limit Copy to* configuration.<br>***To filter the records to be copied:***<br>a. Click **Add** under the *Filter Copied Rows* box.<br>b. From **Source Column** select the column to which the filter should be applied.<br>c. From **Operator** select the SQL operator (for example LIKE or <>).<br>d. From **Filter Value** enter the value to be compared to the *Source Column* using the *Operator*. For example Age >= 18<br>The filters are compiled into an SQL WHERE clause. Multiple filters are combined with an AND operator. A row is copied across when every filter condition is matched. |
| Limit Copy to | By default, a maximum of 100,000 records are copied.<br>**Caution:** This is a protective mechanism. A Cloud database may have tens of millions of records. It is prudent to guard against copy operations that take days to complete and generate unwarranted data storage costs to the Cloud provider. |

**Note:** The copy operation will run as a Toad Background Process. After you click OK, the Copy Table window will remain open until copying is complete.

## Considerations on Copy Table

### When Copying to and from Azure Table Services

Tables in Azure have two columns of data type identifier: a partition key and a row key. Take care in selecting the partition and row keys when copying to and from Azure tables. The partition key is a coarse grain value that groups the rows into discrete subsets of the data that Azure can balance across multiple nodes. (HBase and SimpleDB partition the data automatically but Azure requires the user do it manually.) The row key uniquely identifies each row within the partition. As a good practice the row key should uniquely identify the row within the table.

When copying a table from Azure to a non-Azure target, and the target requires the table have one identifier, change the partition key data type away from identifier. This leaves the row key as the only column data type identifier. The table can now be copied if the row key is unique. If the row key is not unique then identify another column in the Azure table that is unique to be data type identifier.

When copying from a non-Azure table to an Azure target, a column should be selected as the partition key. If a partition key is not selected then an empty string is assigned as the partition key for each row. The two gigabyte limit that Azure enforces for each partition could be exceeded - depending upon the volume of data copied.

To set a partition key:

a. Select the column to use as the partition key. It should divide the table's data into reasonably balanced groups. Suitable candidates are geographical states or regions and departments.

b. For the partition key - set the data type to identifier.

c. For the partition key - select Not Null. A value is required in each row.

d. Move the column identified for the partition key to the top of the list. Use the arrows to the right of the column grid. Toad uses the first identifier column as the partition key and the second as the row key.

### When Copying from an RDBMS to a Cloud Database

Before a table can be copied from an RDBMS to a cloud database, one column of the RDBMS table must be identified as that which has a unique row key. If no such column exists, then a possible workaround is to create a view in the source database that manufactures a unique row key by concatenating two or more columns. This view can be used as the source of the copy operation.

### When Copying to an Oracle Database

When copying to an Oracle database ensure the target remote data object and field names are in UPPERCASE. The Data Hub receives column metadata from Oracle in uppercase, regardless of the letter case the metadata was created in.

### When Copying from One ODBC Database to Another

Temporary Error: *4200 Cannot modify row with no identity fields* may occur when copying a table with no primary key from one ODBC database to another. Although it is possible to create an ODBC table with no primary key it is not good practice to do so.

## Considerations when working with SQL tables

### Network Data Transfer

A query returning a large number of records from a cloud database will take a long time to execute. The aim of the Quest Data Hub is to provide on response times as the worst possible case, but the latency transferring each row must be factored into this calculation.

The Data Hub will communicate filter criteria to the remote database wherever possible to minimize the size of the result set at the data source. This will not help in cases where filtering is

not specified in the SQL, or the volume of data is significant in terms of how long it takes to transfer over the Internet.

### Indexing

The work unit of the cloud world is the map-reduce job, which typically processes an entire file or table each time it runs. The map-reduce paradigm relies upon the data being spread over a large number of nodes to achieve massively parallel processing.

Cloud databases are still very rudimentary in terms of indexing non-identifier fields because they are not required in the map-reduce world. Toad for Cloud Databases uses some strategies to improve the performance of unindexed queries, but there is no performance substitute for indexes in the source data.

Amazon SimpleDB has the best support. It provides a singleton index on every attribute in a domain. Compound filter conditions are supported, but a sub-query is executed for each attribute in the filter and all result sets are subsequently joined according to the logic of the filter. Obviously this strategy is not efficient at dealing with large data volumes where the filtered attribute has relatively few values. Examples of such attributes are gender or state.

### Using count(*) Queries

This is an area where cloud or NoSQL databases conceptually differ from RDBMS. A count(*) query is fairly common in the relational world but quite inappropriate in the cloud world for two reasons. The first is the nature of the problem that cloud databases resolve, specifically the large volumes of data involved in a typical cloud database operation. If counting the rows of a table with five million rows can be slow in an RDBMS, then imagine the cost of counting tens or hundreds of millions in a cloud database.

The second reason pertains to implementation of cloud databases. A count(*) query against a large table in the relational world is usually qualified by a where clause filter that is supported by an index. The index allows the query to execute efficiently despite the large amount of data. Lack of index support in cloud databases can mean the equivalent of a full table scan just to count a handful of rows. Combined with the large data volumes inherent in cloud computing, counting rows is not something that should be treated as a trivial operation.

### Using Views for ODBC Performance

One way to solve the problem of high network overheads is to use views when connecting to a relational database through ODBC. Consider the following query that finds the number of customers serviced by each sales rep:

```
select e.employee_id, e.name, count(*)

from customer c

inner join employee e on e.employee_id = c.sales_rep_id

group by e.employee_id, e.name
```

This query would be perfectly fine in a relational database, provided there was an index on the employee's ID column. The execution plan of this query would be to do a full table scan of the customer table and join each row to the employee table. This query took around half a second when run directly in a SQL Server 2008 database with 5,000 customers spread over 20 sales reps.

Now reconsider this query where that same SQL Server database is mapped in Toad using an ODBC connection. In this case, the Data Hub will need to copy every row from the customer table over the network from the data source, and then make a separate single ODBC fetch from the employee table for each customer row. This latter set-up took around 10 seconds, which will amount to serious time with a large number of rows.

The solution is to create a view in the source database and map that as a table in Toad. The view performs the query in the source relational database without copying the rows across the network.

The next release of Toad will provide a SQL pass through mechanism whereby in-line views – i.e. an inner SELECT statement embedded in the FROM clause of the outer SELECT statement – will be passed straight through to the target data source. This will remove the need to create such a view temporarily for an ad-hoc SQL statement.

### ODBC Table: Error on UPDATE or DELETE with WHERE

With certain ODBC drivers on an ODBC table, executing an SQL UPDATE or DELETE statement with a WHERE clause may return an error. The drivers are attempting to perform the statement while a SELECT cursor is still open on the same connection.

### Oracle ODBC: Table Cannot be Queried. Table or View Does not Exist

Connect directly to the schemas which contain the tables you want to work with or create a synonym for the tables in the schema to which you connect. Even if you see tables in other schemas, you may have problems working with them.

# Appendix: Contact Quest

## Contact Quest Support

Quest Support is available to customers who have a trial version of a Quest product or who have purchased a Quest product and have a valid maintenance contract. Quest Support provides unlimited 24x7 access to SupportLink, our self-service portal. Visit SupportLink at http://support.quest.com.

From SupportLink, you can do the following:

- Retrieve thousands of solutions from our online Knowledgebase
- Download the latest releases and service packs
- Create, update and review Support cases

View the Global Support Guide for a detailed explanation of support programs, online services, contact information, policies and procedures. The guide is available at: http://support.quest.com.

### Toad Community

Get the latest product information, find helpful resources, and join a discussion with the Toad for Cloud Databases team and other community members. Join the Toad for Cloud Databases community at http://toadforcloud.com/index.jspa.

## Contact Quest Software

| | |
|---|---|
| Email | info@quest.com |
| Mail | Quest Software, Inc.<br>World Headquarters<br>5 Polaris Way<br>Aliso Viejo, CA 92656<br>USA |
| Web site | www.quest.com |

See our web site for regional and international office information.

## About Quest Software

Quest Software simplifies and reduces the cost of managing IT for more than 100,000 customers worldwide. Our innovative solutions make solving the toughest IT management problems easier,

enabling customers to save time and money across physical, virtual and cloud environments. For more information about Quest go to www.quest.com.

# Index