SYBASE®

Getting Started — the Aleri Studio

Sybase Aleri Streaming Platform 3.1

DOCUMENT ID: DC01293-01-0311-01

LAST REVISED: June, 2010

Copyright © 2010 Sybase, Inc.

This publication pertains to Sybase software and to any subsequent release until otherwise indicated in new editions or technical notes. Information in this document is subject to change without notice. The software described herein is furnished under a license agreement, and it may be used or copied only in accordance with the terms of that agreement.

To order additional documents, U.S. and Canadian customers should call Customer Fulfillment at (800) 685-8225, fax (617) 229-9845.

Customers in other countries with a U.S. license agreement may contact Customer Fulfillment via the above fax number. All other international customers should contact their Sybase subsidiary or local distributor. Upgrades are provided only at regularly scheduled software release dates. No part of this publication may be reproduced, transmitted, or translated in any form or by any means, electronic, mechanical, manual, optical, or otherwise, without the prior written permission of Sybase, Inc.

Sybase trademarks can be viewed at http://www.sybase.com/detail?id=1011207. Sybase and the marks listed are trademarks of Sybase, Inc. ® indicates registration in the United States of America.

DB2, IBM and Websphere are registered trademarks of International Business Machines Corporation.

Eclipse is a trademark of Eclipse Foundation, Inc.

Excel, Internet Explorer, Microsoft, ODBC, SQL Server, Visual C++, and Windows are trademarks or registered trademarks of Microsoft Corp.

Intel is a registered trademark of Intel Corporation.

JDBC, Solaris, Sun and Sun Microsystems are trademarks or registered trademarks of Sun Microsystems or its subsidiaries in the U.S. and other countries.

Kerberos is a trademark of the Massachusetts Institute of Technology.

Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Netezza is a registered trademark of Netezza Corporation in the United States and/or other countries.

Oracle and Java are registered trademarks of Oracle and/or its affiliates.

SPARC is a registered trademark of SPARC International, Inc. Products bearing SPARC trademarks are based on an architecture developed by Sun Microsystems, Inc.

Teradata is a registered trademark of Teradata Corporation and/or its affiliates in the U.S. and other countries.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Group Ltd.

All other company and product names mentioned may be trademarks of the respective companies with which they are associated.

Use, duplication, or disclosure by the government is subject to the restrictions set forth in subparagraph (c)(1)(ii) of DFARS 52.227-7013 for the DOD and as set forth in FAR 52.227-19(a)-(d) for civilian

agencies.

Sybase, Inc., One Sybase Drive, Dublin, CA 94568.

Table of Contents

About This Guide

1. Related Documents

This guide is part of a set. The following list briefly describes each document in the set.

Product Overview	Introduces the Aleri Streaming Platform and related Aleri products.
Getting Started - the Aleri Studio	Provides the necessary information to start using the Aleri Studio for defining data models.
Release Bulletin	Describes the features, known issues and limitations of the latest Aleri Streaming Platform release.
Installation Guide	Provides instructions for installing and configuring the Streaming Processor and Aleri Studio, which collectively are called the Aleri Streaming Platform.
Authoring Guide	Provides detailed information about creating a data model in the Aleri Studio. Since this is a comprehensive guide, you should read the *Introduction to Data Modeling and the Aleri Studio*. first.
Authoring Reference	Provides detailed information about creating a data model for the Aleri Streaming Platform.
Guide to Programming Interfaces	Provides instructions and reference information for developers who want to use Aleri programming interfaces to create their own applications to work with the Aleri Streaming Platform.
	These interfaces include:
	• the Publish/Subscribe (Pub/Sub) Application Programming Interface (API) for Java
	• the Pub/Sub API for C++
	• the Pub/Sub API for .NET
	• a proprietary Command & Control interface
	• an on-demand SQL query interface
Utilities Guide	Collects usage information (similar to UNIX® man pages) for all Aleri Streaming Platform command line tools.
Administrators Guide	Provides instructions for specific administrative tasks related to the Aleri Streaming Platform.
Introduction to Data Modeling and the Aleri Studio	Walks you through the process of building and testing an Aleri data model using the Aleri Studio.
SPLASH Tutorial	Introduces the SPLASH programming language and illustrates its capabilities through a series of examples.
Frequently Asked Questions	Answers some frequently asked questions about the Aleri Streaming Platform.

Chapter 1. Introduction

1.1. Who Should Use This Guide

This guide provides the necessary information to start using the Aleri Studio for defining data models in the Sybase Aleri Streaming Platform. You'll want to refer to the other guides in the documentation set for more detailed information.

Most users prefer to get started with the Aleri Studio, but data models can also be defined directly in AleriML or Aleri SQL, with the text editor of your choice. You can start the Sybase Aleri Streaming Platform and run from the command line when using AleriML or Aleri SQL. For users that want to take that approach, see the *Authoring Reference Manual*.

1.2. Two Authoring Choices

Sybase's Complex Event Processing (CEP) software - the Sybase Aleri Streaming Platform - loads a data model on startup. The data model defines the input streams, the dataflow, and the operations to produce derived streams.

You have two choices for creating a data model in the Aleri Studio:

Visual Authoring | This is the starting default preferred by most new users. The model is created by dropping streams onto a diagram and directing the flow of data by adding connections between streams.

Aleri SQL | Aleri SQL provides a fast and efficient way of defining a model for more advanced users.
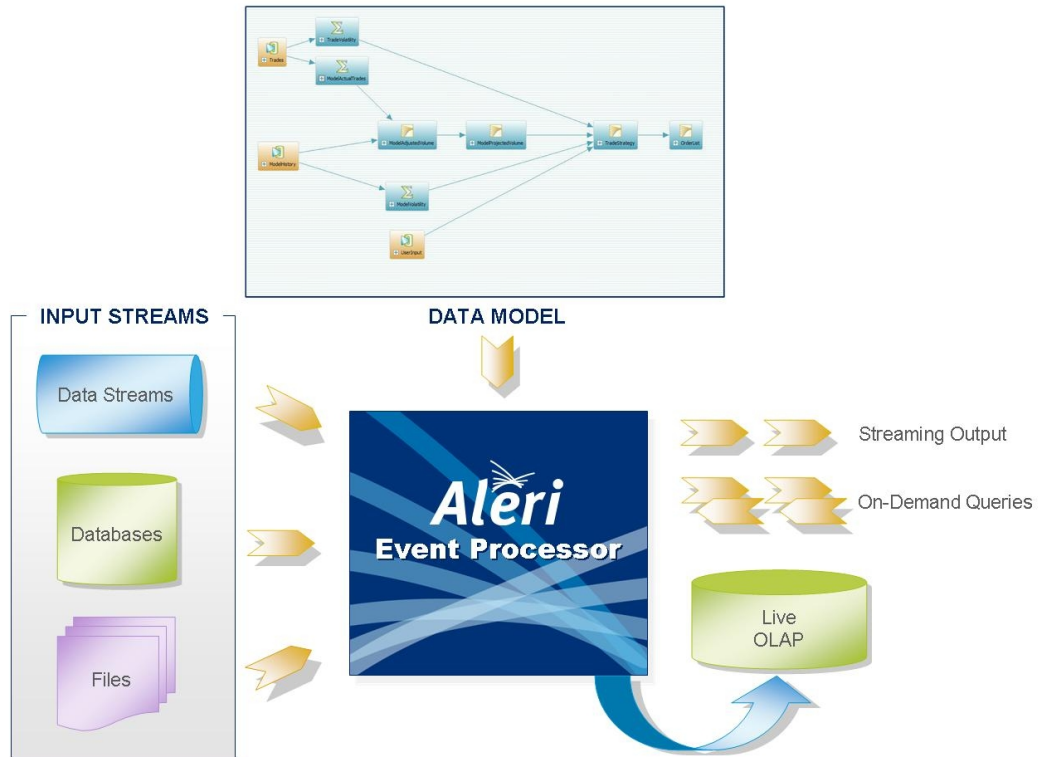
Chapter 2. Visual Authoring

You'll want to start by either creating a new data model or opening an existing data model. Do this from the **File** menu.

2.1. Key Concepts

The Sybase Aleri Streaming Platform loads a data model on startup. The data model contains information about the data sources, data formats and the processing logic to be applied.

A data model consists of one or more source streams and one or more derived streams. All data (streaming, periodically updated, or static) enters the model via a source stream. There can be any number of source streams.

Derived streams are produced by applying one or more operations to an input stream. An input stream can be a source stream or another derived stream. Data flows from source streams to and through any number of derived streams. There are different types of derived streams, each of which has various capabilities with the types of operations it can apply to the incoming data. The flow of data is directed by adding connectors that appear on the diagram as arrows.

Streams in Sybase also store data as incoming events, either adding rows to the table or updating rows in the table. The columns in the table are derived from fields in the event.

Data Locations are external data sources or data destinations. The following are examples of Data Locations: a message bus, a socket, a database, a file or a market data system. Source streams are typically configured with an Input Connection connecting them to a Data Location. Derived streams can be configured with an Output Connection that directs the results to a Data Location that is a destination.

2.2. A Tour of the Workspace
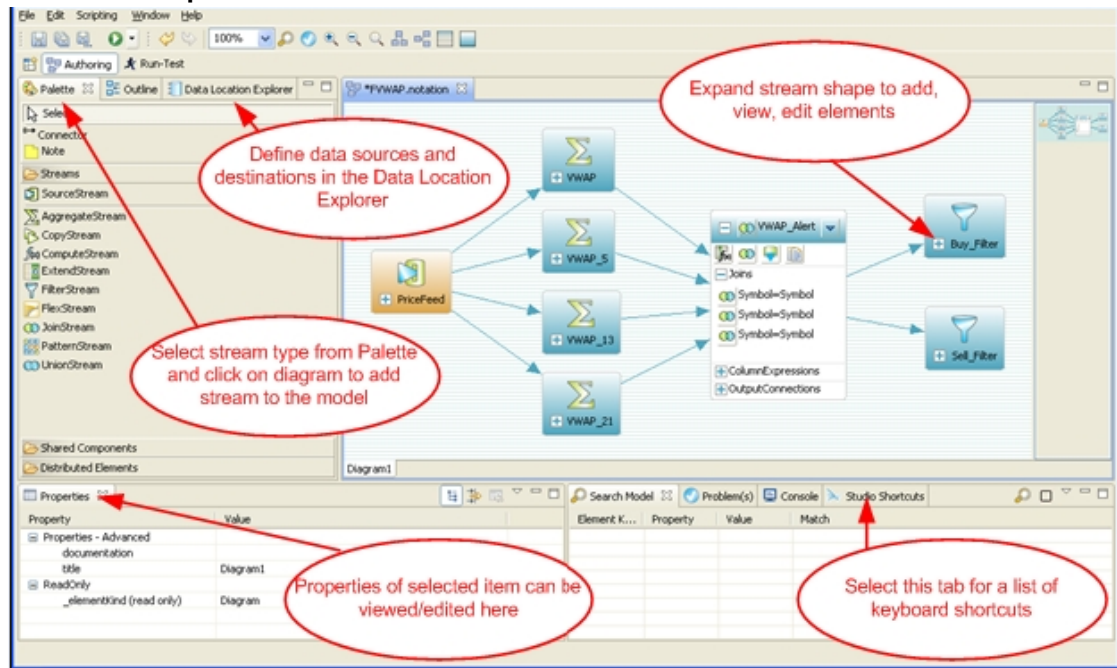
In the center is the diagram editor. This is where you create, view and edit the data model.

On the left is the **Palette**, which lists all of the stream types. To add a stream to the diagram, click on the desired stream type in the **Palette**, then move the cursor to the diagram and click again.

Tucked behind the **Palette** are two important views: the **Outline** and the **Data Location Explorer** (see the tabs along the top of this pane).

The **Data Location Explorer** defines new data locations as well as views and configures existing locations. The **Templates** folder includes templates for various types of Data Locations. The quickest way to define a new Data Location is to start with a template and then set the properties for the specific Data Location.

In the lower left is the **Properties** view. Here you can view and edit the properties of the item currently selected in the diagram.

The **Studio Shortcuts** view in the lower right lists keyboard shortcuts available in the Aleri Studio.

2.3. Data Locations, Input Connections, and Output Connections

Start by defining one or more Data Locations to provide input data to the model. Do this from the **Data Location Explorer** where all data locations are listed. You can create a new Data Location by clicking on **Create DataLocation**, which will bring up the Edit Parameters dialogue.

2.3.1. Data Sources

You can use the Data Discovery feature to import the data format and create a connection if the location exists and supports data discovery. To use the Data Discovery feature, select the Data Location in the explorer and click the Discover tool in the toolbar of the **Data Location Explorer**. If the location is

available, the available data sets (streams, files, tables - depending on location type) will be shown in the Explorer.

2.4. Adding Source Stream

To create a source stream connected to an input data set, select the desired data set in the **Data Location Explorer** and drag it onto the diagram. A new source stream will be created, with an input connection already defined. The column definition of the source stream will also be defined, based on the format of the data source.

2.4.1. Tips

If a particular data location isn't currently available, you can build the model and connect it to the data location later. You can either add a definition for the Data Location at this point or after the model is built. For these Data Locations, you will want to add a source stream to the model manually, by clicking on the stream type in the **Palette**, moving to the desired location on the diagram and clicking again. You should then add the column definitions to the source stream, defining the name and datatype of each column. Later you can add an Input Connection to the source stream to connect it to the Data Location.

To preview the data available from a connection, select the connection from the stream in verbose shape, right click to bring up the context menu and Show->View Sample Data.

If you are connecting to a Data Location that has lots of columns and you only want a subset, rather than dragging the data location onto the diagram to create a source stream, you can create a new source stream from the Palette. You can add an Input Connection to the stream, then right click on the new Input Connection in the stream shape and select Create Columns from Connection.

2.5. Adding Derived Streams

Once you have added some or all of your source streams, you can begin to add derived streams to the model. Click on the desired stream type in the **Palette**, move to the desired location on the diagram and click again add that stream to the model.

There are nine types of Derived Streams:

Aggregate	Groups records according to a common value in one or more columns and produces a single summary record for each group. Takes a single input stream.
Filter	Filters records according to a defined filter expression. Only records on the input stream that cause the filter expression to be 'true' are included in the output. Takes a single input stream.
Compute	Computes output records from input records. The fields in the output record are computed according to an expression defined for each column. Takes a single input stream.
Extend	Is just like a Compute Stream, but does not require an expression for every column in the output stream. Columns in the input stream are passed through unmodified and any new columns are added according to the column expressions defined for the Extend Stream. Takes a single input stream.
Pattern	Watches for patterns of events and produces one or more output records whenever a defined pattern is detected. Takes any number of input streams. Patterns can be defined for a single stream or across multiple input streams.
Join	Joins two or more streams to produce a single output stream. Just like a database join, output records are produced by matching records from the input streams according to common values in selected columns.
Union	Merges two or more streams into a single stream. Input streams must have the same format (column definitions).
Copy	Produces a copy of the input stream. Primarily used for clustered models to provide a local copy of an input stream. Can also be used to effect a self-join.

Flex   Is a programmable stream. Takes any number of input streams. Events on an input stream invoke a method written in the SPLASH scripting language.

Then click on the Connector tool in the **Palette**, click on the stream that will provide input to the new stream, and then click on the new stream. This will add an arrow to the diagram, showing the flow of data.

Tip

If you hold down the **Shift** key when selecting the connector from the palette, you can create multiple connections without selecting the connector each time. When finished adding connections, press the **Esc** key or click on the **Select** tool on the **Palette**.

2.6. Configuring Streams

Most properties of a stream can be set directly in the expanded shape on the diagram. Click on the icon to expand the shape. Once expanded, the shape can be closed by clicking on the icon.
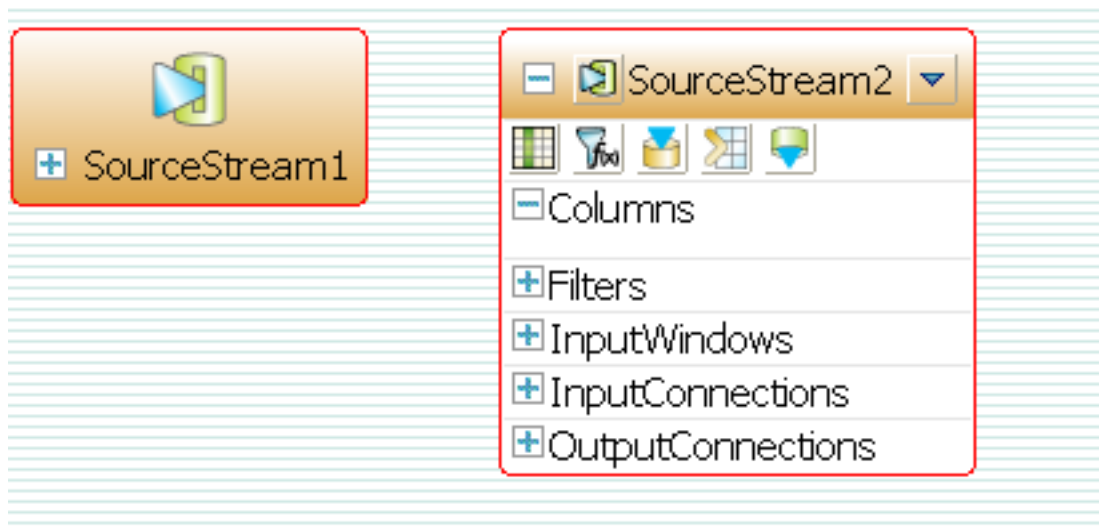
An expanded stream has three main areas: a colored title bar with a pull-down menu in the upper right corner (marked with a triangle pointing down), a toolbar of buttons, and a set of compartments that can be expanded or hidden by clicking on the and icons. The toolbar buttons add components to the compartments.

Source streams have a Columns compartment that defines the names and types of the columns and key columns, a Filters compartment that sets a possible filter expression, and more advanced compartments labeled InputWindows, InputConnections, and OutputConnections.

The compartments vary within derived streams. All have an OutputConnections compartment. Derived streams have an InputWindows compartment except for Join, Flex, and Pattern streams, some have ColumnExpressions compartments (Compute, Extend, Join) whereas others have Columns compartments (Flex and Pattern) for when the names and types of columns cannot be deduced. Copy, Filter, and Union streams have a Read Only Columns compartment that displays the columns inherited from their input streams.

2.7. Adding Output Connections

There are two ways to add output connections:

In verbose shape (mode), click on the tool-button (Add Out Connection), then select the stream. After that, right click to bring up the context menu, choose Create New Child->"Out Connection.

In Outline View, select the stream name and right click to bring up the context menu and choose Create New Child->Out Connection.

Chapter 3. SQL Authoring

SQL authoring is designed for users who are familiar with SQL and want the higher productivity that can be achieved with it compared to visual authoring. However, the Aleri Studio only provides limited assistance for authoring in Aleri SQL. New users are advised to start with Visual Authoring to get familiar with data modeling concepts.

3.1. Key Concepts

Aleri SQL follows the ANSI SQL standard, with some extensions for real-time data streams. Tables represent streams or sets of data from external sources. Materialized Views represent derived streams. A data model consists of one or more Create Table statements and one or more Create Materialized View statements.

See the *Authoring Reference Manual* for a complete definition of the Aleri SQL syntax.

3.2. The Workspace

Data modeling using Aleri SQL is done from the Text Editor Perspective. This perspective simply provides a text editor. Use **Ctrl**-**space** for input assistance. Use the **Check Syntax** button on the toolbar to check for syntax errors before attempting to run the model.

3.3. Running a Model

Use the **Run** button in the toolbar to run the model. A temporary diagram of the model will be created and displayed in the [Visual] Authoring Perspective.

3.4. Files Created

The Sybase Aleri Streaming Platform expects models defined in AleriML. Models written in Aleri SQL are compiled into AleriML before they are run. The AleriML model is contained in a file with the same root name as the SQL file, but with an extension of .xml. In addition, when running an SQL model from the Aleri Studio, a temporary file with the extension of .notation is created in order to display a visual diagram of the model.

Any changes made to the visual diagram will *NOT* be reflected in the SQL model, but changes are applied to the temporary AleriML model that was compiled.

Chapter 4. Running and Testing Your Model

4.1. Quick Start: the Run Button

Press the **Run** button on the toolbar to run the model. This starts the Streaming Processor and loads the active model that is open in the Aleri Studio. (Start-up options can be configured from the drop-down menu available to the right of the **Run** button on the toolbar).

If the model loads successfully, you will see the following message in the console in the lower right of the Studio window:

```
        Server is Alive and Ready_____
```

If errors in the model are found, you will see error messages in the console.

4.2. The Run-Test Perspective

Click on the **Run-Test** button in the Toolbar perspective to access the tools used to test, debug and fine tune a model. Tools available in this perspective include:

Streamviewer	Displays the output of individual streams.
Monitor	Shows queue size, and CPU utilization of each stream in the model.
Debugger	Allows you to set breakpoints in the execution of the model.
Input Tools	Provides the capabilities to record and playback, upload data from files, manually create input records, issue commands to the server and run ad-hoc queries against the server.

4.2.1. Streamviewer

Probably one of the most useful tools for testing a data model. Click on the **Streamviewer** tab in the Run-Test perspective. Click the **Get Stream List** button to see a list of all streams in the model. Double click on a stream to open a live view of the data in that stream.

Chapter 5. AleriRT for Microsoft Excel®

AleriRT for Microsoft Excel® is an add-in that lets you subscribe to and publish real-time data from the Sybase Aleri Streaming Platform. It is not packaged with the Sybase Aleri Streaming Platform; for instructions to download and install it, contact your Sybase representative. You can subscribe to one or more streams and optionally filter and project that data by providing a custom SQL statement. You can publish data automatically whenever a cell in a block changes or manually after selecting a range of cells.

5.1. Installing

To install AleriRT for Microsoft Excel®, run the **setup.exe** utility located in the `IN‐STALL_DIR\drivers\aleriRt\win32` folder. Follow the screen instructions to install the .Net FrameWork 2.0, Windows Installer and a patch for the .Net FrameWork 2.0, if needed. Next, follow the screen instructions for installation.

Afterward, open Microsoft Excel® and verify that the **AleriRT** button is displayed in the toolbar. If it isn't displayed, click the **>>** button to expand the toolbar to display all buttons. When the button is displayed, activate the add-in by opening the Microsoft Excel® program and clicking on **Tools**, then **Add-Ins**. In the screen that appears, click on the **Automation** button and then in the next screen, choose **AleriRT.RTP** from the list box and click on the **OK** button. The **AleriRT.RTP** add-in will now be visible in the first screen with a check mark against it. Now click on the **OK** button.

See the *Installation Guide* for more details.