Esper Reference Documentation

Version: 0.9.1

Table of Contents

Preface

Analyzing and reacting to information in real-time oftentimes requires the development of custom applications. Typically these applications must obtain the data to analyze, filter data, derive information and then indicate this information through some form of presentation or communication. Data may arrive with high frequency requiring high throughput processing. And applications may need to be flexible and react to changes in requirements while the data is processed. Esper is an event stream processor that aims to enable a short development cycle from inception to production for these types of applications.

Esper is a 100% Java component that can be embedded in Java applications. It allows push and pull of data via its subscription and pull API. Esper can be extended by building custom views, functions, windows etc.

1. Read Section 1.1, “Introduction to CEP and event stream analysis” if you are new to CEP and ESP (complex event processing, event stream processing)
2. Read Section 5.1, “Event Pattern Overview” for an overview over event patterns
3. Read Section 6.1, “EQL Introduction” for an introduction to event stream processing via EQL
4. Then glance over the examples Section 9.1, “Examples Overview”

Chapter 1. Technology Overview

1.1. Introduction to CEP and event stream analysis

The Esper engine has been developed to address the requirements of applications that analyze and react to events. Some typical examples of applications are:

- Business process management and automation (process monitoring, BAM, reporting exceptions)
- Finance (algorithmic trading, fraud detection, risk management)
- Network and application monitoring (intrusion detection, SLA monitoring)
- Sensor network applications (RFID reading, scheduling and control of fabrication lines, air traffic)

What these applications have in common is the requirement to process events (or messages) in real-time or near real-time. This is sometimes referred to as complex event processing (CEP) and event stream analysis. Key considerations for these types of applications are throughput, latency and the complexity of the logic required.

- High throughput - applications that process large volumes of messages (between 1,000 to 100k messages per second)
- Low latency - applications that react in real-time to conditions that occur (from a few milliseconds to a few seconds)
- Complex computations - applications that detect patterns among events (event correlation), filter events, aggregate time or length windows of events, join event streams, trigger based on absence of events etc.

The Esper engine was designed to make it easier to build and extend CEP applications.

1.2. CEP and relational databases

Relational databases and the standard query language (SQL) are designed for applications in which most data is fairly static and complex queries are less frequent. Also, most databases store all data on disks (except for in-memory databases) and are therefore optimized for disk access.

To retrieve data from a database an application must issue a query. If an application need the data 10 times per second it must fire the query 10 times per second. This does not scale well to hundreds or thousands of queries per second.

Database triggers can be used to fire in response to database update events. However database triggers tend to be slow and often cannot easily perform complex condition checking and implement logic to react.

In-memory databases may be better suited to CEP applications then traditional relational database as they generally have good query performance. Yet they are not optimized to provide immediate, real-time query results required for CEP and event stream analysis.

1.3. The Esper engine for CEP

The Esper engine works a bit like a database turned upside-down. Instead of storing the data and running queries against stored data, the Esper engine allows applications to store queries and run the data through. Response from the Esper engine is real-time when conditions occur that match queries. The execution model is thus continuous rather then only when a query is submitted.

Esper provides two principal methods or mechanisms to process events: event patterns and event stream queries.

Esper offers an event pattern language to specify expression-based event pattern matching. Underlying the pattern matching engine is a state machine implementation. This method of event processing matches expected sequences of presence or absence of events or combinations of events. It includes time-based correlation of events.

Esper also offers event stream queries that address the event stream analysis requirements of CEP applications. Event stream queries provide the windows, aggregation, joining and analysis functions for use with streams of events. These queries are following the EQL syntax. EQL has been design for similarity with the SQL query language but differs from SQL in its use of views rather then tables. Views represent the different operations needed to structure data in an event stream and to derive data from an event stream.

Esper provides these two methods as alternatives through the same API.

Chapter 2. Architecture

2.1. Overview

A (very) high-level view of the architecture: TODO

2.2. Building and Testing

The Esper code base consists of about 300 source code and 270 unit test (as of release 0.7.0) or test support classes, excluding examples. After build there are over 500 unit test methods that are automatically run to verify the build. Some of the unit tests assert against performance data taken during the test. These tests are designed to run on a single 2.8 GHz Pentium 4 processor with 512MB memory.

Esper requires the following 3rd-party libraries:

- ANTLR is the parser generator used for parsing and parse tree walking of the pattern and EQL syntax. Credit goes to Terence Parr at http://www.antlr.org. The ANTLR license is in the lib directory. The library is required for compile-time only.
- CGLIB is the code generation library for fast method calls. This open source software is under the Apache license. The Apache 2.0 license is in the lib directory.
- LOG4J and Apache commons logging are logging components. This open source software is under the Apache license. The Apache 2.0 license is in the lib directory.
- BeanUtils is a JavaBean manipulation library. This open source software is under the Apache license. The Apache 2.0 license is in the lib directory.
- JUnit is a great unit testing framework. Its license has also been placed in the lib directory. The library is required for build-time only.

Chapter 3. Configuration

Esper engine configuration is entirely optional. Esper has a very small number of configuration parameters that can be used to simplify event pattern and EQL statements, and to tune the engine behavior to specific requirements. The Esper engine works out-of-the-box without configuration.

3.1. Programmatic configuration

An instance of `net.esper.client.Configuration` represents all configuration parameters. The `Configuration` is used to build an (immutable) `EPServiceProvider`, which provides the administrative and runtime interfaces for an Esper engine instance.

You may obtain a `Configuration` instance by instantiating it directly and adding or setting values on it. The `Configuration` instance is then passed to `EPServiceProviderManager` to obtain a configured Esper engine.

```
Configuration configuration = new Configuration();
configuration.addEventTypeAlias("PriceLimit", PriceLimit.class.getName());
configuration.addEventTypeAlias("StockTick", StockTick.class.getName());

EPServiceProvider epService = EPServiceProviderManager.getProvider("MyEngine", configuration);
```

Note that `Configuration` is meant only as an initialization-time object. The Esper engine represented by an `EPServiceProvider` is immutable and does not retain any association back to the `Configuration`.

3.2. Configuration via XML file

An alternative approach to configuration is to specify a configuration in an XML file.

The default name for the XML configuration file is `esper.cfg.xml`. Esper reads this file from the root of the `CLASSPATH` as an application resource via the `configure` method.

```
Configuration configuration = new Configuration();
configuration.configure();
```

The `Configuration` class can read the XML configuration file from other sources as well. The `configure` method accepts `URL, File and String` filename parameters.

```
Configuration configuration = new Configuration();
configuration.configure("myengine.esper.cfg.xml");
```

3.3. XML Configuration file

Here is an example configuration file. The schema for the configuration file can be found in the `etc` folder and is named `esper-configuration-1-0`.

```
<?xml version="1.0" encoding="UTF-8"?>
<esper-configuration xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
           xsi:noNamespaceSchemaLocation="esper-configuration-1-0.xsd">
        <event-type alias="StockTick" class="net.esper.example.stockticker.event.StockTick"/>
        <event-type alias="PriceLimit" class="net.esper.example.stockticker.event.PriceLimit"/>
</esper-configuration>
```

3.4. Configuration items

3.4.1. Event type alias to Java class mapping

This configuration item can be set to allow event pattern statements and EQL statements to use an event type alias rather then the fully qualified Java class name. Interfaces and abstract classes are also supported as event types.

```
every StockTick(symbol='IBM')"
// via configuration equivalent to
every net.esper.example.stockticker.event.StockTick(symbol='IBM')
```

Chapter 4. API Reference

4.1. API Overview

Esper has 2 primary interfaces that this section outlines: The administrative interface and the runtime interface.

Use Esper's administrative interface to create event patterns and EQL statements as discussed in Section 5.1, “Event Pattern Overview” and Section 6.1, “EQL Introduction”.

Use Esper's runtime interface to send events into the engine, emit events and get statistics for an engine instance.

The JavaDoc documentation is also a great source for API information.

4.2. Engine Instances

Each instance of an Esper engine is completely independent of other engine instances and has its own administrative and runtime interface.

An instance of the Esper engine is obtained via static methods on the `EPServiceProviderManager` class. The `getDefaultProvider` method and the `getProvider(String URI)` methods return an instance of the Esper engine. The latter can be used to obtain multiple instances of the engine for different URI values. The `EPServiceProviderManager` determines if the URI matches all prior URI values and returns the same engine instance for the same URI value. If the URI has not been seen before, a new Engine instance is created.

The code snipped below gets the default instance Esper engine. Subsequent calls to get the default engine instance return the same instance.

```
EPServiceProvider epService = EPServiceProviderManager.getDefaultProvider();
```

This code snippet gets an Esper engine for URI `RFIDProcessor1`. Subsequent calls to get an engine with the same URI return the same instance.

```
EPServiceProvider epService = EPServiceProviderManager.getProvider("RFIDProcessor1");
```

An existing Esper engine instance can be reset via the `initialize` method on the `EPServiceProvider` instance. This stops and removes all statements in the Engine.

4.3. The Administrative Interface

Create event patterns or EQL statements via the administrative interface `EPAdministrator`.

This code snippet gets an Esper engine then creates an event pattern and an EQL statement.

```
EPServiceProvider epService = EPServiceProviderManager.getDefaultProvider();
EPAdministrator admin = epService.getEPAdministrator();
EPStatement 10secRecurTrigger = admin.createPattern("every timer:at(*, *, *, *, *, */10)");
EPStatement weightedAvgView = admin.createEQL(
    "select * from MarketDataBean(symbol='IBM').win:time(60).stat:weighted_avg('price', 'volume')");
```

The `createPattern` and `createEQL` methods return `EPStatement` instances. Statements are automatically started and active when created. A statement can also be stopped and started again via the `stop` and `start` methods shown in the code snippet below.

```
weightedAvgView.stop();
weightedAvgView.start();
```

We can subscribe to updates posted by a statement via the `addListener` and `removeListener` methods the `EPStatement` statement. We need to provide an implementation of the `UpdateListener` interface to the statement.

```
UpdateListener myListener = new MyUpdateListener(); // MyUpdateListener implements UpdateListener
weightedAvgView.addListener(myListener);
```

EQL statements and event patterns publish old data and new data to registered `UpdateListener` listeners. Old data published by views constists of the events representing the prior values of derived data held by the statement. New data published by views is the events representing the new values of derived data held by the statement.

Subscribing to events posted by a statement is following a push model. The engine pushes data to listeners when events are received that cause data to change or patterns to match. Alternatively, statements can also serve up data in a pull model via the `iterator` method. This can come in handy if we are not interested in all new updates, but only want to perform a frequent poll for the latest data. For example, an event pattern that fires every 5 seconds could be used to pull data from an EQL statement. The code snippet below demonstrates some pull code.

```
Iterator<EventBean> eventIter = weightedAvgView.iterator();
for (EventBean event : eventIter) {
   // .. do something ..
}
```

This is a second example:

```
double averagePrice = (Double) eqlStatement.iterator().next().get("average");
```

4.4. The Runtime Interface

The `EPRuntime` interface is used to send events for processing into an Esper engine, and to emit Events from an engine instance to the outside world.

The below code snippet shows how to send events to the engine.

```
EPServiceProvider epService = EPServiceProviderManager.getDefaultProvider();
EPRuntime runtime = epService.getEPRuntime();

// Send an example event containing stock market data
runtime.sendEvent(new MarketDataBean('IBM', 75.0));
```

Another important method in the runtime interface is the `route` method. This method is designed for use by `UpdateListener` implementations that need to send events into an engine instance.

The `emit` and `addEmittedListener` methods can be used to emit events from a runtime to a registered set of one or more emitted event listeners. Events are emitted on an event channel identified by a name. Listeners are implementations of the `EmittedListener` interface. Listeners can specify a channel to listen to and thus only receive events posted to that channel. Listeners can also supply no channel name and thus receive emitted

events posted on any channel. Channels are uniquely identified by a string channel name.

4.5. Event Class Requirements

An event is an immutable record of a past occurence of an action or state change. An event can have a set of event properties that supply information about the event. An event also has an event class.

In Esper, events are object instances that expose event properties through JavaBean-style getter methods. Events classes or interfaces do not have to be fully compliant to the JavaBean specification; however for the Esper engine to obtain event properties, the required JavaBean getter methods must be present.

Esper supports JavaBean-style event classes that extend a superclass or implement one or more interfaces. Also, Esper event pattern and EQL statements can refer to Java interface classes and abstract classes.

Classes that represent events should be made immutable. As events are recordings of a state change or action that occured in the past, the relevant event properties should not be changable. However this is not a hard requirement and the Esper engine accepts events that are mutable as well.

4.5.1. Event Property Types

The set of possible property types supported by a JavaBean and by Esper can be broken into below categories -- some of which are supported by the standard JavaBeans specification, and some of which are uniquely supported by Esper:

- *Simple* properties have a single value that may be retrieved. The underlying property type might be a Java language primitive (such as int, a simple object (such as a java.lang.String), or a more complex object whose class is defined either by the Java language, by the application, or by a class library included with the application.
- *Indexed* - An indexed property stores an ordered collection of objects (all of the same type) that can be individually accessed by an integer-valued, non-negative index (or subscript). Alternatively, the entire set of values may be retrieved using an array.
- *Mapped* - As an extension to standard JavaBeans APIs, Esper considers any property that accepts a String-valued key a mapped property.
- *Nested* - A nested property is a property that lives within another Java object which itself is a property of an event.

Assume there is an EmployeeEvent event class as shown below. The mapped and indexed properties in this example return Java objects but could also return Java language primitive types (such as int or String). The Address object and Employee objects can themselves have properties that are nested within them, such as a streetName in the Address object or a name of the employee in the Employee object.

```
public class EmployeeEvent {
        public String getFirstName();
        public Address getAddress(String type);
        public Employee getSubordinate(int index);
        public Employee[] getAllSubordinates();
}
```

Simple event properties require a getter-method that returns the property value. In this example, the `getFirstName` getter method returns the `firstName` event property of type String.

Indexed event properties require either one of the following getter-methods. A method that takes an integer-type key value and returns the property value, such as the `getSubordinate` method. Or a method that returns an array-type such as the `getSubordinates` getter method, which returns an array of Employee. In an EQL or

event pattern statement, indexed properties are accessed via the `property[index]` syntax.

Mapped event properties require a getter-method that takes a String-typed key value and returns the property value, such as the `getAddress` method. In an EQL or event pattern statement, mapped properties are accessed via the `property('key')` syntax.

Nested event properties require a getter-method that returns the nesting object. The `getAddress` and `getSubordinate` methods are mapped and indexed properties that return a nesting object. In an EQL or event pattern statement, nested properties are accessed via the `property.nestedProperty` syntax.

All event pattern and EQL statements allow the use of indexed, mapped and nested properties (or a combination of these) anywhere where one or more event property names are expected. The below example shows different combinations of indexed, mapped and nested properties in filters of event pattern expressions.

```
every EmployeeEvent(firstName='myName')
every EmployeeEvent(address('home').streetName='Park Avenue')
every EmployeeEvent(subordinate[0].name='anotherName')
every EmployeeEvent(allSubordinates[1].name='thatName')
every EmployeeEvent(subordinate[0].address('home').streetName='Water Street')
```

Similarly, the syntax can be used in EQL statements in all places where an event property name is expected, such as in select lists, where-clauses or join criteria.

```
select firstName, address('work'), subordinate[0].name, subordinate[1].name
from EmployeeEvent
where address('work').streetName = 'Park Ave'
```

4.6. Time-Keeping Events

Special events are provided that can be used to control the time-keeping of an engine instance. There are two models for an engine to keep track of time. Internal clocking is when the engine instance relies on the `java.util.Timer` class for time tick events. External clocking can be used to supply time ticks to the engine. The latter is useful for testing time-based event sequences or for synchronizing the engine with an external time source.

By default, the Esper engine uses internal time ticks. This behavior can be changed by sending a timer control event to the engine as shown below.

```
EPServiceProvider epService = EPServiceProviderManager.getDefaultProvider();
EPRuntime runtime = epService.getEPRuntime();
// switch to external clocking
runtime.sendEvent(new TimerControlEvent(TimerControlEvent.ClockType.CLOCK_EXTERNAL));

// send a time tick
long timeInMillis = System.currentTimeMillis(); // Or get the time somewhere else
runtime.sendEvent(new CurrentTimeEvent(timeInMillis));
```

4.7. Events Received from the Engine

The Esper engine posts events to registered `UpdateListener` instances ('push' method for receiving events). For many statements events can also be pulled from statements via the `iterator` method. Both pull and push supply `EventBean` instances representing the events generated by the engine or events supplied to the engine. Each `EventBean` instance represents an event, with each event being either an artificial event, composite event or an event supplied to the engine via its runtime interface.

The `getEventType` method supplies an event's event type information represented by an `EventType` instance. The `EventType` supplies event property names and types as well as information about the underlying object to the event.

The engine may generate artificial events that contain information derived from event streams. A typical example for artificial events is the events posted for a statement to calculate univariate statistics on an event property. The below example shows such a statement and queries the generated events for an average value.

```
// Derive univariate statistics on price for the last 100 market data events
String viewExpr = "select * from MarketDataBean(symbol='IBM').win:length(100).stat:uni('price')";
EPStatement priceStatsView = epService.getEPAdministrator().createEQL(viewExpr);
priceStatsView.addListener(testListener);
```

```
// Example listener code
public class MyUpdateListener implements UpdateListener
{
    public void update(EventBean[] newData, EventBean[] oldData)
    {
                // Interrogate events
                System.out.println("new average price=" + newData[0].get("average");
        }
}
```

Composite events are events that aggregate one or more other events. Composite events are typically created by the engine for statements that join two event streams, and for event patterns in which the causal events are retained and reported in a composite event. The example below shows such an event pattern.

```
// Look for a pattern where AEvent follows BEvent
String pattern = "a=AEvent -> b=BEvent";
EPStatement stmt = epService.getEPAdministrator().createPattern(pattern);
stmt.addListener(testListener);
```

```
// Example listener code
public class MyUpdateListener implements UpdateListener
{
    public void update(EventBean[] newData, EventBean[] oldData)
    {
                System.out.println("a event=" + newData[0].get("a").getUnderlying());
                System.out.println("b event=" + newData[0].get("b").getUnderlying());
        }
}
```

Chapter 5. Event Pattern Reference

5.1. Event Pattern Overview

Event patterns match when an event or multiple events occur that match the pattern's definition. Patterns can also be time-based.

Pattern expressions can consist of filter expressions combined with pattern operators. Expressions can contain further nested pattern expressions by including the nested expression(s) in `()` round brackets.

There are 5 types of operators:

1. Operators that control pattern finder creation and termination: `every`
2. Logical operators: `and`, `or`, `not`
3. Temporal operators that operate on event order: `->` (followed-by)
4. Guards are where-conditions that filter out events and cause termination of the pattern finder. Examples are `timer:within`.
5. Observers observe time events as well as other events. Examples are `timer:interval` and `timer:at`.

5.2. How to use Patterns

5.2.1. Pattern Syntax

The `createPattern` method on the `EPAdministrator` administrative interface creates pattern statements for the given pattern expression string.

This is an example pattern expression that matches on every `ServiceMeasurement` events in which the value of the `latency` event property is over 20 seconds, and on every`ServiceMeasurement` events in which the `success` property is false. Either one or the other condition must be true for this pattern to match.

```
every (spike=ServiceMeasurement(latency>20000) or error=ServiceMeasurement(success=false))
```

The Java code to create this trigger is below.

```
EPAdministrator admin = EPServiceProviderManager.getDefaultProvider().getEPAdministrator();

String eventName = ServiceMeasurement.class.getName();

EPStatement myTrigger = admin.createPattern(
  "every (spike=" + eventName + "(latency>20000) or error=" + eventName + "(success=false))");
```

The pattern expression starts with an `every` operator to indicate that the pattern should fire for every matching events and not just the first matching event. Within the `every` operator in round brackets is a nested pattern expression using the `or` operator. The left hand of the `or` operator is a filter expression that filters for events with a high latency value. The right hand of the operator contains a filter expression that filters for events with error status. Filter expressions are explained in Section 5.3, “Filter Expressions”.

5.2.2. Subscribing to Pattern Events

When a pattern fires it publishes one or more events to any listeners to the pattern statement. The listener inter-

face is the `net.esper.client.UpdateListener` interface.

The example below shows an anonymous implementation of the `net.esper.client.UpdateListener` interface. We add the anonymous listener implementation to the `myPattern` statement created earlier. The listener code simply extracts the underlying event class.

```
myPattern.addListener(new UpdateListener()
{
        public void update(EventBean[] newEvents, EventBean[] oldEvents)
        {
                ServiceMeasurement spike = (ServiceMeasurement) newEvents[0].get("spike");
                ServiceMeasurement error = (ServiceMeasurement) newEvents[0].get("error");
                ... // either spike or error can be null, depending on which occured
                ... // add more logic here
        }
});
```

Listeners receive an array of `EventBean` instances in the `newEvents` parameter. There is one `EventBean` instance passed to the listener for each combination of events that matches the pattern expression. At least one `EventBean` instance is always passed to the listener.

The properties of each `EventBean` instance contain the underlying events that caused the pattern to fire, if events have been named in the filter expression via the `name=eventType` syntax. The property name is thus the name supplied in the pattern expression, while the property type is the type of the underlying class, in this example `ServiceMeasurement`.

5.2.3. Pulling Data from Patterns

Data can also be pulled from pattern statements via the `iterator()` method. If the pattern had fired at least once, then the iterator returns the last event for which it fired. The `hasNext()` method can be used to determine if the pattern had fired.

```
if (myPattern.iterator().hasNext())
{
        ServiceMeasurement event = (ServiceMeasurement) view.iterator().next().get("alert");
    ... // some more code here to process the event
}
else
{
    ... // no matching events at this time
}
```

5.3. Filter Expressions

This chapter outines how to filter events based on their properties.

The simplest form of filter is a filter for events of a given type without any conditions on the event property values. This filter matches any event of that type regardless of the event's properties. The example below is such a filter. Note that this event pattern would stop firing as soon as the first RfidEvent is encountered.

```
com.mypackage.myevents.RfidEvent
```

To make the event pattern fire for every RfidEvent and not just the first event, use the `every` keyword.

```
every com.mypackage.myevents.RfidEvent
```

The example above specifies the fully-qualified Java class name as the event type. Via configuration, the event pattern above can be simplified by using the alias that has been defined for the event type. Interfaces and abstract classes are also supported as event types.

```
every RfidEvent
```

Interfaces and superclasses are also supported as event types. In the below example `IRfidReadable` is an interface class.

```
every org.myorg.rfid.IRfidReadable
```

The filtering criteria to filter for events with certain event property values are placed within parenthesis after the event type name.

```
mypackage.RfidEvent(category="Perishable")
```

The supported filter operators are

- equals `=`
- comparison operators `< , > , >=, <=`
- ranges use the keyword `in` and round `(...)` or square brackets `[]`

Ranges come in the following 4 varieties. The use of round `()` or square `[]` bracket dictates whether an endpoint is included or excluded.

- Open ranges that contain neither endpoint `(low:high)`
- Closed ranges that contain both endpoints `[low:high]`
- Half-open ranges that contain the low endpoint but not the high endpoint `[low:high)`
- Half-closed ranges that contain the high endpoint but not the low endpoint `(low:high]`

Filter criteria are listed in a comma-separated format. In the example below we look for `RfidEvent` events with a `grade` property between 1 and 2 (endpoints included), a `price` less then 1, and a category of "Perishable".

```
mypackage.RfidEvent(category="Perishable", price<1.00, grade in [1:2])
```

Filter criteria can also refer to events machting prior named events in the same expression. Below pattern is an example in which the pattern matches once for every RfidEvent that is preceded by an RfidEvent with the same item id.

```
every A=mypackage.RfidEvent -> B=mypackage.RfidEvent(itemId=A.itemId)
```

The syntax shown above allows filter criteria to reference prior results by specifying the event name and event property. This syntax can be used with all filter operators.

Some limitations of filters are:

- Range and comparison operators require the event property to be of a numeric type.
- Null values in filter criteria are currently not allowed.
- Filter criteria can list the same event property only once.
- Events that have null values for event properties listed in the filter criteria do not match the criteria.

5.4. Pattern Operators

5.4.1. Every

The `every` operator indicates that the pattern expression should restart when the pattern matches. Without the `every` operator the pattern expressions matcher stops when the pattern matches once.

Thus the `every` operator works like a factory for the pattern expression contained within. When the pattern expression within it fires and thus quits checking for events, the `every` causes the start of a new pattern matcher listening for more occurances of the same event or set of events.

Every time a pattern expression within an `every` operator turns true a new active pattern matcher is started looking for more event(s) or timing conditions that match the pattern expression. If the `every` operator is not specified for an expression, the expression stops after the first match was found.

This pattern fires when encountering event A and then stops looking.

```
A
```

This pattern keeps firing when encountering event A, and doesn't stop looking.

```
every A
```

Let's consider an example event sequence as follows.

A_1 B_1 C_1 B_2 A_2 D_1 A_3 B_3 E_1 A_4 F_1 B_4

Table 5.1. 'Every' operator examples

Example	Description
`every ( A -> B )`	Detect event A followed by event B. At the time when B occurs the pattern matches, then the pattern matcher restarts and looks for event A again. 1. Matches on B_1 for combination {A_1, B_1} 2. Matches on B_3 for combination {A_2, B_3} 3. Matches on B_4 for combination {A_4, B_4}
`every A -> B`	The pattern fires for every event A followed by an event B. 1. Matches on B_1 for combination {A_1, B_1} 2. Matches on B_3 for combination {A_2, B_3} and {A_3, B_3} 3. Matches on B_4 for combination {A_4, B_4}
`A -> every B`	The pattern fires for an event A followed by every event B. 1. Matches on B_1 for combination {A_1, B_1}. 2. Matches on B_2 for combination {A_1, B_2}. 3. Matches on B_3 for combination {A_1, B_3} 4. Matches on B_4 for combination {A_1, B_4}
`every A -> every B`	The pattern fires for every event A followed by every event B. 1. Matches on B_1 for combination {A_1, B_1}. 2. Matches on B_2 for combination {A_1, B_2}.

Example	Description
	3. Matches on B_3 for combination $\{A_1, B_3\}$ and $\{A_2, B_3\}$ and $\{A_3, B_3\}$ 4. Matches on B_4 for combination $\{A_1, B_4\}$ and $\{A_2, B_4\}$ and $\{A_3, B_4\}$ and $\{A_4, B_4\}$

The examples show that it is possible that a pattern fires for multiple combinations of events that match a pattern expression. Each combination is posted as an `EventBean` instance to the `update` method in the `UpdateListener` implementation.

5.4.2. And

Similar to the Java && operator the `and` operator requires both nested pattern expressions to turn true before the whole expression turns true (a join pattern).

Pattern matches when both event A and event B are found.

```
A and B
```

Pattern matches on any sequence A followed by B and C followed by D, or C followed by D and A followed by B

```
(A -> B) and (C -> D)
```

5.4.3. Or

Similar to the Java "||" operator the `or` operator requires either one of the expressions to turn true before the whole expression turns true.

Look for either event A or event B. As always, A and B can itself be nested expressions as well.

```
A or B
```

Detect all stock ticks that are either above or below a threshold.

```
every (StockTick(symbol='IBM', price < 100) or StockTick(symbol='IBM', price > 105)
```

5.4.4. Not

The `not` operator negates the truth value of an expression. Pattern expressions prefixed with `not` are automatically defaulted to true.

This pattern matches only when an event A is encountered followed by event B but only if no event C was encountered before event B.

```
( A -> B ) and not C
```

5.4.5. Followed-by

The followed by `->` operator specifies that first the left hand expression must turn true and only then is the right hand expression evaluated for matching events.

Look for event A and if encountered, look for event B. As always, A and B can itself be nested event pattern expressions.

```
A -> B
```

This is a pattern that fires when 2 status events indicating an error occur one after the other.

```
StatusEvent(status='ERROR') -> StatusEvent(status='ERROR')
```

5.5. Guards

5.5.1. timer:within

The `timer:within` guard acts like a stopwatch. If the associated pattern expression does not turn true within the specified time period it is stopped and permanently false.

This pattern fires for all A events that arrive within 5 seconds.

```
every A where timer:within (5000)
```

This pattern matches for any A or B events in the next 5 seconds.

```
( A or B ) where timer:within (5000)
```

This pattern matches for any 2 errors that happen 10 seconds within each other.

```
every (StatusEvent(status='ERROR') -> StatusEvent(status='ERROR') where timer:within (10000))
```

5.6. Pattern Observers

5.6.1. timer:interval

The `timer:interval` observer takes a wait time in milliseconds and waits for the defined time before the truth value of the observer turns true.

After event A arrived wait 10 seconds then indicate that the pattern matches.

```
A -> timer:interval(10000)
```

The pattern below fires every 20 seconds.

```
every timer:interval(20000)
```

5.6.2. timer:at

The `timer:at` observer is similar in function to the Unix “crontab” command. At a specified time the expression turns true. The `at` operator can also be made to pattern match at regular intervals by using an `every` operator in front of the `timer:at` operator.

The syntax is: `timer:at (minutes, hours, days of month, months, days of week [, seconds]).`

The value for seconds is optional. Each element allows wildcard `*` values. Ranges can be specified by means of lower bounds then a colon ':' then the upper bound. The division operator `*/x` can be used to specify that every x_{th} value is valid. Combinations of these operators can be used by placing these into square brackets([]).

This expression pattern matches every 5 minutes past the hour.

```
every timer:at(5, *, *, *, *)
```

The below at operator pattern matches every 15 minutes from 8am to 5pm on even numbered days of the month as well as on the first day of the month.

```
timer:at (*/15, 8:17, [*/2, 1], *, *)
```

Chapter 6. EQL Reference

6.1. EQL Introduction

EQL statements are used to derive and aggregate information from one or more streams of events, and to join event streams. This section outlines EQL syntax. It also outlines the built-in views, which are the building blocks for deriving and aggregating information from event streams.

EQL is similar to SQL in its use of the `select` clause and the `where` clause. Where EQL differs most from SQL is in the use of tables. EQL replaces tables with the concept of event streams.

EQL statements contain definitions of one or more views. Similar to tables in an SQL statement, views define the data available for querying and filtering. Some views represent windows over a stream of events. Other views derive statistics from event properties, group events or handle unique event property values. Views can be staggered onto each other to build a chain of views. The Esper engine makes sure that views are reused among EQL statements for efficiency.

The built-in set of views is:

1. Views that represent moving event windows: `win:length`, `win:time`, `win:time_batch`, `win:ext_time`, `ext:sort_window`
2. Views for aggregation: `std:unique`, `std:groupby`, `std:lastevent`
3. Views that derive statistics: `std:size`, `stat:uni`, `stat:linest`, `stat:correl`, `stat:weighted_avg`, `stat:multidim_stat`

Esper can be extended by plugging-in custom developed views.

6.2. EQL Syntax

EQL queries are created and stored in the engine, and publish results as events are received by the engine or timer events occur that match the criteria specified in the query. Events can also be pulled from running EQL queries.

The select clause in an EQL query specifies the event properties or events to retrieve. The `from` clause in an EQL query specifies the event stream definitions and stream names to use. The `where` clause in an EQL query specifies search conditions that specify which event or event combination to search for. For example, the following statement returns the average price for IBM stock ticks in the last 30 seconds if the average hit 75 or more.

```
select average from StockTick(symbol='IBM').win:time(30).stat:uni('price') where average >= 75;
```

EQL queries follow the below syntax. EQL queries can be simple queries or more complex queries. A simple select contains only a select clause and a single stream definition. Complex EQL queries can be build that feature a more elaborate select list utilizing expressions, may join multiple streams or may contain a where clause that with search conditions.

```
select select_list
from stream_def [as name] [, stream_def [as name]] [,...]
[where search_conditions]
[group by grouping_expression_list]
[having grouping_search_conditions]
[output output_specification]
```

```
```

6.3. Choosing Event Properties And Events: the *Select* Clause

The select clause is required in all EQL statements. The select clause can be used to select all properties via the wildcard `*`, or to specify a list of event properties and expressions. The select clause defines the event type (event property names and types) of the resulting events published by the statement, or pulled from the statement.

6.3.1. Choosing all event properties: select *

The syntax for selecting all event properties in a stream is:

```
select * from stream_def
```

The following statement selects all univariate statistics properties for the last 30 seconds of IBM stock ticks for price.

```
select * from StockTick(symbol='IBM').win:time(30).stat:uni('price')
```

In a join statement, using the `select *` syntax selects event properties that contain the events representing the joined streams themselves.

6.3.2. Choosing specific event properties

To chose the particular event properties to return:

```
select event_property [, event_property] [, ...] from stream_def
```

The following statement selects the count and standard deviation properties for the last 100 events of IBM stock ticks for volume.

```
select count, stdev from StockTick(symbol='IBM').win:length(100).stat:uni('volume')
```

6.3.3. Expressions

The select clause can contain one or more expressions.

```
select expression [, expression] [, ...] from stream_def
```

The following statement selects the volume multiplied by price for a time batch of the last 30 seconds of stock tick events.

```
select volume * price from StockTick.win:time_batch(30)
```

6.3.4. Renaming event properties

Event properties and expressions can be renamed using below syntax.

```
select [event property | expression] as identifier [, ...]
```

The following statement selects volume multiplied by price and specifies the name *volPrice* for the event property.

```
select volume * price as volPrice from StockTick.win:length(100)
```

6.4. Specifying Event Streams : the *From* Clause

The `from` clause is required in all EQL statements. It specifies one or more event streams. Each event stream can optionally be given a name by means of the `as` syntax.

```
from stream_def [as name] [, stream_def [as name]] [, ...]
```

The event stream definition *stream_def* as shown in the syntax above consists of an event type, an optional filter property list and an optional list of views that derive data from a stream must be supplied. The syntax for an event stream definition is as below:

```
event_type ( [filter_criteria] ) [.view_spec] [.view_spec] [...]
```

The following EQL statement selects all event properties for the last 100 events of IBM stock ticks for volume. In the example, the event type is the fully qualified Java class name `org.esper.example.StockTick`. The expression filters for events where the property `symbol` has a value of "IBM". The optional view specifications for deriving data from the StockTick events are a length window and a view for computing statistics on volume. The name for the event stream is "volumeStats".

```
select * from org.esper.example.StockTick(symbol='IBM').win:length(100).stat:uni('volume') as volumeSt
```

Instead of the fully-qualified Java class name any other event name can be mapped via Configuration to a Java class, making the resulting statement more readable.

```
select * from StockTick(symbol='IBM').win:length(100).stat:uni('volume') as volumeStats
```

6.4.1. Specifying an event type

In the example above the event type was `org.esper.example.StockTick`. The event type is simply the fully qualified Java class name. Interfaces and abstract classes are also supported. Alternatively, via configuration an alias for an event type can be defined and used instead of the fully qualified class name. The below example shows one way to obtain the fully qualified class name of a given Java class `StockTick`.

```
String eventName = StockTick.class.getName();
String stmt = "from " + eventName + ".win:length(100)"
```

6.4.2. Specifying event filter criteria

Filter criteria follow the same syntax as outlined in the event pattern section on filters; see Section 5.3, “Filter Expressions”. Filter criteria operators are: `=, < , > , >=, <=`. Ranges use the `in` keyword and round `(...)` or square brackets `[]`.

Esper filters out events in an event stream as defined by filter criteria before it sends events to subsequent views. Thus, compared to search conditions in a where-clause, filter criteria remove unneeded events early.

The below example is a filter criteria list that removes events based on category, price and grade.

```
from mypackage.RfidEvent(category="Perishable", price<1.00, grade in [1, 2])
```

6.4.3. Specifying views

Views are used to derive or aggregate data. Views can be staggered onto each other. The section below outlines the views available and plug-in of custom views.

Views can optionally take one or parameters. These parameters can consist of primitive constants such as String, boolean or numeric types. String arrays are also supported as a view parameter type.

Views can optionally take one or parameters. These parameters can consist of primitive constants such as String, boolean or numeric types. String arrays are also supported as a view parameter type.

The below example uses the car location event. It specifies an empty list of filter criteria by adding a empty round brackets `()` after the event type. The first view `"std:groupby('carId')"` groups car location events by car id. The second view `"win:length(4)"` keeps a length window of the 4 last events, with one length window for each car id. The next view `"std:groupby({'expressway', 'direction', 'segment'})"` groups each event by it's expressway, direction and segment property values. Again, the grouping is done for each car id considering the last 4 events only. The last view `"std:size()"` is used to report the number of events. Thus the below example reports the number of events per car id and per expressway, direction and segment considering the last 4 events for each car id only. The `"as accSegment"` syntax assigns the name accSegment to the resulting event stream.

```
String carLocEvent = CarLocEvent.class.getName();
String joinStatement = "select * from " +
        carLocEvent + ".std:groupby('carId').win:length(4).std:groupby({'expressway', 'direction', 'se
        carLocEvent + ".win:time(30).std:unique('carId') as curCarSeg" +
                " where accSeg.size >= 4" +
                "   and accSeg.expressway = curCarSeg.expressway" +
                "   and accSeg.direction = curCarSeg.direction" +
                "   and (" +
                                "(accSeg.direction=0 " +
                                " and curCarSeg.segment < accSeg.segment" +
                                " and curCarSeg.segment > accSeg.segment - 5)" +
                        " or " +
                                "(accSeg.direction=1 " +
                                " and curCarSeg.segment > accSeg.segment" +
                                " and curCarSeg.segment < accSeg.segment + 5)" +
                        ")";
```

6.5. Specifying Search Conditions: the *Where* Clause

The where clause is an optional clause in EQL statements. Via the where clause event streams can be joined and events can be filtered.

Comparison operators `=, < , > , >=, <=, !=, <>, is null, is not null` and logical combinations via `and` and `or` are supported in the where clause. The where clause can also introduce join conditions as outlined in Section 6.10, “Joining Event Streams”. Where-clauses can also contain expressions. Some examples are listed below.

```
...where fraud.severity = 5 and amount > 500
...where (orderItem.orderId is null) or (orderItem.class != 10)
...where (orderItem.orderId = null) or (orderItem.class <> 10)
...where itemCount / packageCount > 10
```

6.6. Aggregates and grouping: the *Group-by* Clause and the *Having* Clause

6.6.1. Using aggregate functions

The aggregate functions are `sum, avg, count, max, min, median, stddev, avedev`. You can use aggregate functions to calculate and summarize data from event properties. For example, to find out the total price for all stock tick events in the last 30 seconds, type:

```
select sum(price) from StockTickEvent.win:time(30)
```

Here is the syntax for aggregate functions:

```
aggregate_function( [all | distinct] expression)
```

You can apply aggregate functions to all events in an event stream window or other view, or to one or more groups of events. From each set of events to which an aggregate function is applied, Esper generates a single value.

`Expression` is usually an event property name. However it can also be a constant, function, or any combination of event property names, constants, and functions connected by arithmetic operators.

For example, to find out the average price for all stock tick events in the last 30 seconds if the price was doubled:

```
select avg(price * 2) from StockTickEvent.win:time(30)
```

You can use the optional keyword `distinct` with all aggregate functions to eliminate duplicate values before the aggregate function is applied. The optional keyword `all` which performs the operation on all events is the default.

The syntax of the aggregation functions and the results they produce are shown in below table.

Table 6.1. Syntax and results of aggregate functions

Aggregate Function	Result
sum([all|distinct] *expression*)	Totals the (distinct) values in the expression, returning a value of `long, double, float or integer` type depending on the expression
avg([all|distinct] *expression*)	Average of the (distinct) values in the expression, returning a value of `double` type
count([all|distinct] *expression*)	Number of the (distinct) non-null values in the expression, returning a value of `long` type
count(*)	Number of events, returning a value of `long` type
max([all|distinct] *expression*)	Highest (distinct) value in the expression, returning a value of the same

Aggregate Function	Result
	type as the expression itself returns
min([all|distinct] *expression*)	Lowest (distinct) value in the expression, returning a value of the same type as the expression itself returns
median([all|distinct] *expression*)	Median (distinct) value in the expression, returning a value of `double` type
stddev([all|distinct] *expression*)	Standard deviation of the (distinct) values in the expression, returning a value of `double` type
avedev([all|distinct] *expression*)	Mean deviation of the (distinct) values in the expression, returning a value of `double` type

You can use aggregation functions in a `select` clause and in a `having` clause. You cannot use aggregate functions in a `where` clause, but you can use the where clause to restrict the events to which the aggregate is applied. The next query computes the average and sum of the price of stock tick events for the symbol IBM only, for the last 10 stock tick events regardless of their symbol.

```
select 'IBM stats' as title, avg(price) as avgPrice, sum(price) as sumPrice
from StockTickEvent.win:length(10)
where symbol='IBM'
```

In the above example the length window of 10 elements is not affected by the where-clause, i.e. all events enter and leave the length window regardless of their symbol. If we only care about the last 10 IBM events, we need to add filter criteria as below.

```
select 'IBM stats' as title, avg(price) as avgPrice, sum(price) as sumPrice
from StockTickEvent(symbol='IBM').win:length(10)
where symbol='IBM'
```

You can use aggregate functions with any type of event property or expression, with the following exceptions:

1. You can use `sum, avg, median, stddev, avedev` with numeric event properties only

Esper ignores any null values returned by the event property or expression on which the aggregate function is operating, except for the `count(*)` function, which counts null values as well. All aggregate functions return null if the data set contains no events, or if all events in the data set contain only null values for the aggregated expression.

6.6.2. Organizing statement results into groups: the *Group-by* clause

The `group by` clause is optional in all EQL statements. The `group by` clause divides the output of an EQL statement into groups. You can group by one or more event property names, or by the result of computed expressions. When used with aggregate functions, `group by` retrieves the calculations in each subgroup. You can use `group by` without aggregate functions, but generally that can produce confusing results.

For example, the below statement returns the total price per symbol for all stock tick events in the last 30 seconds:

```
select symbol, sum(price) from StockTickEvent.win:time(30) group by symbol
```

The syntax of the group by clause is:

```
group by arregate_free_expression [, arregate_free_expression] [, ...]
```

Esper places the following restrictions on expressions in the `group by` clause:

1. Expressions in the `group by` cannot contain aggregate functions
2. Event properties that are used within aggregate functions in the `select` clause cannot also be used in a `group by` expression

You can list more then one expression in the `group by` clause to nest groups. Once the sets are established with `group by` the aggregation functions are applied. This statement posts the median volume for all stock tick events in the last 30 seconds per symbol and tick data feed. Esper posts one event for each group to statement listeners:

```
select symbol, tickDataFeed, median(volume)
from StockTickEvent.win:time(30)
group by symbol, tickDataFeed
```

In the statement above the event properties in the `select` list (symbol, tickDataFeed) are also listed in the `group by` clause. The statement thus follows the SQL standard which prescribes that non-aggregated event properties in the `select` list must match the `group by` columns.

Esper also supports statements in which one or more event properties in the `select` list are not listed in the `group by` clause. The statement below demonstrates this case. It calculates the standard deviation for the last 30 seconds of stock ticks aggregating by symbol and posting for each event the symbol, tickDataFeed and the standard deviation on price.

```
select symbol, tickDataFeed, stddev(price) from StockTickEvent.win:time(30) group by symbol
```

The above example still aggregates the `price` event property based on the `symbol`, but produces one event per incoming event, not one event per group.

Additionally, Esper supports statements in which one or more event properties in the `group by` clause are not listed in the `select` list. This is an example that calculates the mean deviation per `symbol` and `tickDataFeed` and posts one event per group with `symbol` and mean deviation of price in the generated events. Since tick-DataFeed is not in the posted results, this can potentially be confusing.

```
select symbol, avedev(price)
from StockTickEvent.win:time(30)
group by symbol, tickDataFeed
```

Expressions are also allowed in the `group by` list:

```
select symbol * price, count(*) from StockTickEvent.win:time(30) group by symbol * price
```

If the `group by` expression resulted in a null value, the null value becomes its own group. All null values are aggregated into the same group. If you are using the `count(expression)` aggregate function which does not count null values, the count returns zero if only null values are encountered.

You can use a `where` clause in a statement with `group by`. Events that do not satisfy the conditions in the `where` clause are eliminated before any grouping is done. For example, the statement below posts the number of stock ticks in the last 30 seconds with a volume larger then 100, posting one event per group (symbol).

```
select symbol, count(*) from StockTickEvent.win:time(30) where volume > 100 group by symbol
```

6.6.3. Selecting groups of events: the *Having* clause

Use the `having` clause to pass or reject events defined by the `group-by` clause. The `having` clause sets conditions for the `group by` clause in the same way `where` sets conditions for the `select` clause, except `where` cannot include aggregate functions, while `having` often does.

This statement is an example of a `having` clause with an aggregate function. It posts the total price per symbol for the last 30 seconds of stock tick events for only those symbols in which the total price exceeds 1000. The `having` clause eliminates all symbols where the total price is equal or less then 1000.

```
select symbol, sum(price)
from StockTickEvent.win:time(30)
group by symbol
having sum(price) > 1000
```

To include more then one condition in the `having` clause combine the conditions with `and`, `or` or `not`. This is shown in the statement below which selects only groups with a total price greater then 1000 and an average volume less then 500.

```
select symbol, sum(price), avg(volume)
from StockTickEvent.win:time(30)
group by symbol
having sum(price) > 1000 and avg(volume) < 500
```

Esper places the following restrictions on expressions in the `having` clause:

1. Any expressions that contain aggregate functions must also occur in the `select` clause

A statement with the `having` clause should also have a `group by` clause. If you omit `group-by`, all the events not excluded by the `where` clause return as a single group. In that case `having` acts like a `where` except that `having` can have aggregate functions.

The `having` clause can also be used without `group by` clause as the below example shows. The example below posts events where the price is less then the current running average price of all stock tick events in the last 30 seconds.

```
select symbol, price, avg(price)
from StockTickEvent.win:time(30)
having price < avg(price)
```

6.6.4. How the stream filter, *Where*, *Group By* and *Having* clauses interact

When you include filters, the `where` condition, the `group by` clause and the `having` condition in an EQL statement the sequence in which each clause affects events determines the final result:

1. The event stream's filter condition, if present, dictates which events enter a window (if one is used). The filter discards any events not meeting filter criteria.
2. The `where` clause excludes events that do not meet its search condition.
3. Aggregate functions in the select list calculate summary values for each group.
4. The `having` clause excludes events from the final results that do not meet its search condition.

The following query illustrates the use of filter, `where`, `group by` and `having` clauses in one statement with a

`select` clause containing an aggregate function.

```
select tickDataFeed, stddev(price)
from StockTickEvent(symbol='IBM').win:length(10)
where volume > 1000
group by tickDataFeed
having stddev(price) > 0.8
```

Esper filters events using the filter criteria for the event stream `StockTickEvent`. In the example above only events with symbol IBM enter the length window over the last 10 events, all other events are simply discarded. The `where` clause removes any events posted by the length window (events entering the window and event leaving the window) that do not match the condition of volume greater then 1000. Remaining events are applied to the `stddev` standard deviation aggregate function for each tick data feed as specified in the `group by` clause. Each `tickDataFeed` value generates one event. Esper applies the `having` clause and only lets events pass for `tickDataFeed` groups with a standard deviation of price greater then 0.8.

6.7. Stabilizing and Limiting Output: the *Output* Clause

6.7.1. Output Clause Options

The `output` clause is optional in Esper and is used to stabilize the rate at which events are output. For example, the following statement batches old and new events and outputs them at the end of every 90 second interval.

```
select * from StockTickEvent.win:length(5) output every 90 seconds
```

Here is the syntax for output rate limiting:

```
output [all | last] every number [minutes | seconds | events]
```

The optional `last` keyword specifies to only output the very last event, while the `all` keyword is the default and specifies to output all events in a batch. The batch size can be specified in terms of time or number of events.

The time interval can also be specified in terms of minutes; the following statement is identical to the first one.

```
select * from StockTickEvent.win:length(5) output every 1.5 minutes
```

A second way that output can be stabilized is by batching events until a certain number of events have been collected. The next statement only outputs when either 5 (or more) new or 5 (or more) old events have been batched.

```
select * from StockTickEvent.win:time(30) output every 5 events
```

Additionally, event output can be further modified by the optional last keyword, which causes output of only the last event to arrive into an output batch.

```
select * from StockTickEvent.win:time(30) output last every 5 events
```

6.7.2. Group By, Having and Output clause interaction

The `output` clause interacts in two ways with the `group by` and `having` clauses. First, in the `output every n events` case, the number `n` refers to the number of events arriving into the `group by clause`. That is, if the

`group by` clause outputs only 1 event per group, or if the arriving events don't satisfy the `having` clause, then the actual number of events output by the statement could be fewer than `n`.

Second, the `last` and `all` keywords have special meanings when used in a statement with aggregate functions and the `group by` clause. The `last` keyword specifies that only groups whose aggregate values have been updated with the most recent batch of events should be output. The `all` keyword (the default) specifies that the most recent data for all groups seen so far should be output, whether or not these groups' aggregate values have just been updated.

6.8. Single-row function reference

Single-row functions return a single value for every single event result row generated by your statement. These functions can appear in the `select` clause, in the `where` clause and in the `having` clause.

The below table outlines the single-row functions available.

Table 6.2. Syntax and results of single-row functions

Single-row Function	Result
max(*expression*, *expression*, [, *expression* [,...]])	Returns the highest numeric value among the 2 or more comma-separated expressions.
min(*expression*, *expression*, [, *expression* [,...]])	Returns the lowest numeric value among the 2 or more comma-separated expressions.

An example showing the use of the `min` single-row function is below.

```
select symbol, min(ticks.timestamp, news.timestamp) as minT
from StockTickEvent.win:time(30) as ticks,
     NewsEvent.win:time(30) as news
where ticks.symbol = news.symbol
```

6.9. Build-in views

This chapter outlines the views that are built into Esper.

6.9.1. Window views

Length window

Creates a moving window extending the specified number of elements into the past.

The below example calculates basic univariate statistics for the last 5 stock ticks for symbol IBM.

```
StockTickEvent(symbol='IBM').win:length(5).stat:uni('price')
```

The next example keeps a length window of 10 events of stock trade events, with a separate window for each symbol. The statistics on price is calculated only for the last 10 events for each symbol.

```
StockTickEvent.std:groupby('symbol').win:length(10).stat:uni('price')
```

Time window

The time_window creates a moving time window extending from the specified time interval in seconds into the past based on the system time.

For the IBM stock tick events in the last 1000 milliseconds, calculate statistics on price.

```
StockTickEvent(symbol='IBM').win:time(1).stat:uni('price')
```

Externally-timed window

Similar to the time window this view moving time window extending from the specified time interval in seconds into the past, but based on the millisecond time value supplied by an event property.

This view holds stock tick events of the last 10 seconds based on the timestamp property in `StockTickEvent`.

```
StockTickEvent.win:ext_timed(10, 'timestamp')
```

Time window buffer

This window view buffers events and releases them every specified time interval in one update.

The below example batches events into a 5 second window releasing new batches every 5 seconds. Listeners to updates posted by this view receive updated information only every 5 seconds.

```
StockTickEvent.win:time_batch(5)
```

6.9.2. Standard view set

Unique

The `unique`view is a view that includes only the most recent among events having the same value for the specified field.

The below example creates a view that retains only the last event per symbol.

```
StockTickEvent.std:unique('symbol')
```

Group By

This view groups events into sub-views by the value of the specified field.

This example calculates statistics on price separately for each symbol.

```
StockTickEvent.std:groupby('symbol').stat:uni('price')
```

The next example keeps a length window of 10 events of stock trade events, with a separate window for each symbol. Now the statistics on price is calculated only for the last 10 events for each symbol.

```
StockTickEvent.std:groupby('symbol').win:length(10).stat:uni('price')
```

Size

This view returns the number of elements in view.

This example view reports the number of events within the last 1 minute.

```
StockTickEvent.win:time(60000).std:size()
```

Last

This view exposes the last element of its parent view.

This example view retains statistics calculated on stock tick price for the symbol IBM.

```
StockTickEvent(symbol='IBM').stat:uni('price').std:lastevent()
```

6.9.3. Statistics views

Univariate statistics

This view calculated basic univariate statistics on an event property.

Table 6.3. Univariate statistics derived properties

Property Name	Description
`count`	Number of values
`sum`	Sum of values
`average`	Average of values
`variance`	Variance
`stdev`	Sample standard deviation (square root of variance)
`stdevpa`	Population standard deviation

The below example calculates price statistics on stock tick events for the last 10 events.

```
StockTickEvent.win:length(10).stat:uni('price')
```

Regression

This view calculates regression on two event properties.

Table 6.4. Regression derived properties

Property Name	Description
`slope`	Slope
`yintercept`	Y Intercept

Calculate slope and y-intercept on price and offer for all events in the last 10 seconds.

```
StockTickEvent.win:time(10000).stat:linest('price', 'offer')
```

Correlation

This view calculates the correlation value on two event properties.

Table 6.5. Correlation derived properties

Property Name	Description
`correl`	Correlation between two event properties

Calculate correlation on price and offer over all stock tick events for IBM.

```
StockTickEvent(symbol='IBM').stat:correl('price', 'offer')
```

Weighted average

This view returns the weighted average given a weight field and a field to compute the average for. Syntax: weighted_avg(field, weightField)

Table 6.6. Weighted average derived properties

Property Name	Description
`average`	Weighted average

Views that derive the volume-weighted average price for the last 3 seconds.

```
StockTickEvent(symbol='IBM').win:time(3000).stat:weighted_avg('price', 'volume')
```

Multi-dimensional statistics

This view works similar to the `std:groupby` views in that it groups information by one or more event properties. The view accepts 3 or more parameters: The first parameter to the view defines the univariate statistics values to derive. The second parameter is the property name to derive data from. The remaining parameters supply the event property names to use to derive dimensions.

Table 6.7. Multi-dim derived properties

Property Name	Description
`cube`	The cube following the interface

The example below derives the count, average and standard deviation latency of service measurement events per customer.

```
ServiceMeasurement.stat:multidim_stats({‘count’, ‘average’, ‘stdev’},
    'latency', 'customer')
```

This example derives the average latency of service measurement events per customer, service and error status for events in the last 30 seconds.

```
ServiceMeasurement.win:length(30000).stat:multidim_stats({‘average’},
        'latency', 'customer', 'service', 'status')
```

6.9.4. Extension View Set

Sorted Window View

This view sorts by values in the specified event property and keeps only the top elements up to the given size.

The syntax for this view is : `sort(String propertyName, boolean isDescending, int size) .`

These view can be used to sort on price descending keeping the lowest 10 prices and reporting statistics on price.

```
StockTickEvent.ext:sort('price', true, 10).stat:uni('price'))
```

6.10. Joining Event Streams

Two or more event streams can be part of the `from` clause and thus both streams determine the resulting events. The where-clause lists the join conditions that Esper uses to relate events in the two or more streams.

Each point in time that an event arrives to one of the event streams, the two event streams are joined and output events are produced according to the where-clause.

This example joins 2 event streams. The first event stream consists of fraud warning events for which we keep the last 30 minutes (1800 seconds). The second stream is withdrawal events for which we consider the last 30 seconds. The streams are joined on account number.

```
select fraud.accountNumber as accntNum, fraud.warning as warn, withdraw.amount as amount,
       max(fraud.timestamp, withdraw.timestamp) as timestamp, 'withdrawlFraud' as desc
  from net.esper.example.atm.FraudWarningEvent.win:time(1800) as fraud,
       net.esper.example.atm.WithdrawalEvent.win:time(30) as withdraw
 where fraud.accountNumber = withdraw.accountNumber
```

6.11. Outer Join

Esper supports left outer joins, right outer joins and full outer joins between an unlimited number of event streams.

If the outer join is a left outer join, there will be an output event for each event of the stream on the left-hand side of the clause. For example, in the left outer join shown below we will get output for each event in the stream RfidEvent, even if the event does not match any event in the event stream OrderList.

```
select * from net.esper.example.rfid.RfidEvent.win:time(30) as rfid
       left outer join
       net.esper.example.rfid.OrderList.win:length(10000) as orderlist
     on rfid.itemId = orderList.itemId
```

Similarly, if the join is a Right Outer Join, then there will be an output event for each event of the stream on the

right-hand side of the clause. For example, in the right outer join shown below we will get output for each event in the stream OrderList, even if the event does not match any event in the event stream RfidEvent.

```
select * from net.esper.example.rfid.RfidEvent.win:time(30) as rfid
       right outer join
       net.esper.example.rfid.OrderList.win:length(10000) as orderlist
     on rfid.itemId = orderList.itemId
```

For all types of outer joins, if the join condition is not met, the select list is computed with the event properties of the arrived event while all other event properties are considered to be null.

```
select * from net.esper.example.rfid.RfidEvent.win:time(30) as rfid
       full outer join
       net.esper.example.rfid.OrderList.win:length(10000) as orderlist
     on rfid.itemId = orderList.itemId
```

The last type of outer join is a full outer join. In a full outer join, each point in time that an event arrives to one of the event streams, one or more output events are produced. In the example below, when either an RfidEvent or an OrderList event arrive, one or more output event is produced.

6.12. View Plug-in

This is currently not supported (planned).

Chapter 7. Adapters

This chapter discusses adapters (TODO)

7.1. Adapter

Adapters adapt event executions in the outside world into a format for processing by Esper, and feed events to Esper.

Currently there are no pre-build adapters available for Esper.

Chapter 8. Indicators

8.1. Intro

Indicators are pluggable modules that communicate the results of event stream processing to the external world. Indicators can act as visualizers that present a graphical view of their event inputs. They can also be warning agents (monitors) that send alerts, warnings or other control events to the outside world.

In their implementation indicators can be classes that implement the `UpdateListener` interface and that can thus be attached directly to one or more statements. Indicactors can also be attached to one or more `EPStatement` instances. This makes is possible for indicators to merge data as well as pull data from trigger and statement views.

Indicators may be integration components that plug together with other software, an some indicators will be supplied by Esper. Esper currently only has one indicator module as described below.

8.2. JMX Indicator

The `net.esper.indicator.jmx.JMXLastEventIndicator` displays the last event in a JMX MBean it registers with the MBeanServer obtained via ManagementFactory.getPlatformMBeanServer();

Chapter 9. Examples

9.1. Examples Overview

This chapters outlines the examples that come with Esper in the `eg/src` folder of the distribution. The code for examples can be found in the `net.esper.example` packages.

In order to compile and run the samples please follow the below instructions:

1. Make sure Java 1.5 or greater is installed and the JAVA_HOME environment variable is set.
2. Open a console window and change directory to esper/eg/etc.
3. Run "setenv.bat" (Windows) or "setenv.sh" (Unix) to verify your environment settings.
4. Run "compile.bat" (Windows) or "compile.sh" (Unix) to compile the examples.
5. Now you are ready to run the examples. Some examples require mandatory parameters. Further information to running each example can be found in "readme.txt".
6. Modify the logger logging level in the "log4j.xml" configuration file changing DEBUG to INFO on a class or package level to reduce the volume of text output.

JUnit tests exist for the example code. The JUnit test source code for the examples can be found in the `eg/test` folder. To build and run the example JUnit tests, use the Maven 2 goal `test`. The JUnit test source code can also be helpful in understanding the example and in the use of Esper APIs.

9.2. Transaction 3-Event Challenge

The classes for this example live in package `net.esper.example.transaction`. Run "run_txnsim.bat" (Windows) or "run_txnsim.sh" (Unix) to start the transaction simulator. Please see the readme file in the same folder for build instructions and command line parameters.

9.2.1. The Events

The use case involves tracking three components of a transaction. It‘s important that we use at least three components, since some engines have different performance or coding for only two events per transaction. Each component comes to the engine as an event with the following fields:

- Transaction ID
- Time stamp

In addition, we have the following extra fields:

In event A:

- Customer ID

In event C:

- Supplier ID (the ID of the supplier that the order was filled through)

9.2.2. Combined event

We need to take in events A, B and C and produce a single, combined event with the following fields:

- Transaction ID
- Customer ID
- Time stamp from event A
- Time stamp from event B
- Time stamp from event C

What we‘re doing here is matching the transaction IDs on each event, to form an aggregate event. If all these events were in a relational database, this could be done as a simple SQL join… except that with 10,000 events per second, you will need some serious database hardware to do it.

9.2.3. Real time summary data

Further, we need to produce the following:

- Min,Max,Average total latency from the events (difference in time between A and C) over the past 30 minutes.
- Min,Max,Average latency grouped by (a) customer ID and (b) supplier ID. In other words, metrics on the the latency of the orders coming from each customer and going to each supplier.
- Min,Max,Average latency between events A/B (time stamp of B minus A) and B/C (time stamp of C minus B).

9.2.4. Find problems

We need to detect a transaction that did not make it through all three events. In other words, a transaction with events A or B, but not C. Note that, in this case, what we care about is event C. The lack of events A or B could indicate a failure in the event transport and should be ignored. Although the lack of an event C could also be a transport failure, it merits looking into.

9.2.5. Event generator

To make testing easier, standard and to demonstrate how the example works, the example is including an event generator. The generator generates events for a given number of transactions, using the following rules:

- One in 5,000 transactions will skip event A
- One in 1,000 transactions will skip event B
- One in 10,000 transactions will skip event C.
- Transaction identifiers are randomly generated
- Customer and supplier identifiers are randomly chosen from two lists
- The time stamp on each event is based on the system time. Between events A and B as well as B and C, between 0 and 999 is added to the time. So, we have an expected time difference of around 500 milliseconds between each event
- Events are randomly shuffled as described below

To make things harder, we don‘t want transaction events coming in order. This code ensures that they come completely out of order. To do this, we fill in a bucket with events and, when the bucket is full, we shuffle it. The buckets are sized so that some transactions‘ events will be split between buckets. So, you have a fairly randomized flow of events, representing the worst case from a big, distributed infrastructure.

The generator lets you change the size of the bucket (small, medium, large, larger, largerer). The larger the bucket size, the more events potentially come in between two events in a given transaction and so, the more the performance characteristics like buffers, hashes/indexes and other structures are put to the test as the bucket size increases.

9.3. StockTicker

The StockTicker example comes from the stock trading domain. The example creates event patterns to filter stock tick events based on price and symbol. When a stock tick event is encountered that falls outside the lower or upper price limit, the example simply displays that stock tick event. The price range itself is dynamically created and changed. This is accomplished by an event patterns that searches for another event class, the price limit event.

The classes `net.esper.example.stockticker.event.StockTick` and `PriceLimit` represent our events. The event patterns are created by the class `net.esper.example.stockticker.monitor.StockTickerMonitor`.

Summary:

- Good example to learn the API and get started with event patterns
- Dynamically creates and removes event patterns based on price limit events received
- Simple, highly-performant filter expressions for event properties in the stock tick event such as symbol and price

9.4. MatchMaker

In the MatchMaker example every mobile user has an X and Y location, a set of properties (gender, hair color, age range) and a set of preferences (one for each property) to match. The task of the event patterns created by this example is to detect mobile users that are within proximity given a certain range, and for which the properties match preferences.

The event class representing mobile users is `net.esper.example.matchmaker.event.MobileUserBean`. The `net.esper.example.matchmaker.monitor.MatchMakingMonitor` class contains the patterns for detecing matches.

Summary:

- Dynamically creates and removes event patterns based on mobile user events received
- Uses range matching for X and Y properties of mobile user events

9.5. QualityOfService

This example develops some code for measuring quality-of-service levels such as for a service-level agreement (SLA). A SLA is a contract between 2 parties that defines service constraints such as maximum latency for service operations or error rates.

The example measures and monitors operation latency and error counts per customer and operation. When one of our operations oversteps these constraints, we want to be alerted right away. Additionally, we would like to have some monitoring in place that checks the health of our service and provides some information on how the operations are used.

Some of the constraints we need to check are:

- That the latency (time to finish) of some of the operations is always less then X seconds.
- That the latency average is always less then Y seconds over Z operation invocations.

The `net.esper.example.qos_sla.events.OperationMeasurement` event class with its latency and status properties is the main event used for the SLA analysis. The other event `LatencyLimit` serves to set latency limits on the fly.

The `net.esper.example.qos_sla.monitor.AverageLatencyMonitor` creates an EQL statement that computes latency statistics per customer and operation for the last 100 events. The `DynaLatencySpikeMonitor` uses an event pattern to listen to spikes in latency with dynamically set limits. The `ErrorRateMonitor` uses the timer `'at'` operator in an event pattern that wakes up periodically and polls the error rate within the last 10 minutes. The `ServiceHealthMonitor` simply alerts when 3 errors occur, and the `SpikeAndErrorMonitor` alerts when a fixed latency is overstepped or an error status is reported.

Summary:

- This example combines event patterns with EQL statements for event stream analysis.
- Shows the use of the timer `'at'` operator and followed-by operator `->` in event patterns
- Outlines basic EQL statements
- Shows how to pull data out of EQL statements rather then subscribing to events a statement publishes

9.6. LinearRoad

The Linear Road example is a very incomplete implementation of the Stream Data Management Benchmark [3] by Standford University.

Linear Road simulates a toll system for the motor vehicle expressways of a large metropolitan area. The main event in this example is a car location report which the class `net.esper.example.linearroad.CarLocEvent` represents. Currently the event stream joins are performed by JUnit test classes in the `eg/test` folder. See the `net.esper.example.linearroad.TestAccidentNotify` and the `TestCarSegmentCount` classes. Please consider this a work in progress.

Summary:

- Shows more complex joins between event streams.

9.7. StockTick RSI

The RSI gives you the trend for a stock and for more complete explanation, you can visit the link: http://www.stockcharts.com/education/IndicatorAnalysis/indic_RSI.html.

After a definite number of stock events, or accumulation period, the first RSI is computed. Then for each subsequent stock event, the RSI calculations use the previous period’s Average Gain and Loss to determine the “smoothed RSI”.

Summary:

- Uses a simple event pattern with a filter which feeds a listener that computes the RSI, which publishes events containing the computed RSI.

Chapter 10. References

10.1. Reference List

- Luckham, David. 2002. *The Power of Events.* Addison-Wesley.
- The Stanford Rapide (TM) Project. *http://pavg.stanford.edu/rapide*.
- Arasu, Arvind, et.al.. 2004. Linear Road: A Stream Data Management Benchmark, Stanford University *http://www.cs.brown.edu/research/aurora/Linear_Road_Benchmark_Homepage.html*.