Space Details

Key:	PentahoDoc
Name:	BI Server Documentation - Latest
Description:	Latest version of the Pentaho BI Server
Creator (Creation Date):	admin (Nov 15, 2006)
Last Modifier (Mod. Date):	admin (Nov 27, 2006)

Available Pages

- Subscriptions Users Guide
  - 01. Introduction to Subscriptions
  - 02. Configuring the Subscription Subsystem
  - 03. Subscription Admin
  - 04. User Subscriptions
  - 05. CRON Expressions in Detail

Subscriptions Users Guide

This page last changed on Mar 16, 2007 by gmoran.

Important Version Information

Subscriptions is a feature that has moved from our Professional version into open source. If you are using the open source version of our platform prior to milestone build 1.5.2, then subscriptions will not be available. To get the Subscriptions feature, update your distribution to the latest release.

This document describes how to use the Subscriptions functionality of Pentaho BI Suite Professional Edition. It assumes familiarity with the Pentaho BI Platform and Action Sequence documents. The easiest way to work with a Pentaho BI Server is by installing the Pre-Configured Install (PCI) on your local machine.

Subscriptions allow users to schedule and configure the delivery of content within the platform. So what does that mean? Scheduling refers to the periodic execution of Action Sequences at selected times. Configuration is the ability for a user to set the parameters that will be passed to the Action Sequence at execution time. Delivery is usually via email although this is not a requirement. Content is the output of BI Processes such as reports, email notifications, exceptions etc. All of this will be explained in greater detail - but first, a little review and background.

01. Introduction to Subscriptions

This page last changed on Mar 21, 2007 by mlowery.

Subscriptions allow users to schedule and configure the delivery of content within the platform. So what does that mean? Scheduling refers to the periodic execution of Action Sequences at selected times. Configuration is the ability for a user to set the parameters that will be passed to the Action Sequence at execution time. Delivery is usually via email although this is not a requirement. Content is the output of BI Processes such as reports, email notifications, exceptions etc. All of this will be explained in greater detail - but first, a little review and background.

Roles

The Pentaho BI Platform is a set of components and services that can be configured and deployed many different ways. This can make identifying the people using the system in different capacities confusing. A generic term like **user** can imply different things. A user may be the person using the platform to create components for solutions or the person using that within their own solution or the user that ultimately uses the solution. To make the people and the roles more clear, the following conventions will be used:

users : When referring to users in this document, we mean the end users of the solutions that are created by solution developers. A user task would be selecting a report and possible subscribing to it.

administrator: The person responsible for configuring and monitoring the Pentaho BI Platform server that is serving users. Usually an IS person whose responsibilities are related to keeping the hardware, operating system, application server, networking and database server functioning.

solution developer: The person that creates the Action Sequences, reports, database tables etc. The solution developer will typically build solutions that users will ultimately use.

subscription administrator: The administrator would define and manage schedules and content that can be subscribed to. They will typically use the **Subscription Admin** utility to manage the subscriptions. In many cases the solution developer and subscription administrator may be the same person.

Review of Action Sequences

Activities that the Pentaho BI platform performs are defined by Action Sequence XML documents. Action Sequences define activities such as database queries, report generation, email actions and the order in which they occur. These activities are called Actions. Data can be passed between Actions within an Action Sequence or between executions of other Action Sequences. Action Sequence documents define the inputs that must be available prior to execution, the processes that will be performed (the Actions) and the outputs that will be available to other Action Sequences after execution is complete. The *Creating Pentaho Solutions* guide provides an in depth explanation of Pentaho BI Server architecture and the role Action Sequence documents. It also provides a reference for the XML and the interface to components within the platform. We recommend reading the *Creating Pentaho Solutions* guide prior to reading this document.

Design Studio

The Pentaho Design Studio is used to graphically build and test Action Sequence documents. It provides a place to organize solutions and uses templates and wizards to help automate some of the more common tasks. Being familiar with the Pentaho Design Studio will make building and testing subscription content much easier than doing it all by hand.

CRON Expressions

The subscription subsystem uses the Quartz scheduler to specify when to execute the Action Sequence that creates the content that the user has subscribed to. In order to specify the times, the platform uses the UNIX standard cron expression. The cron expression has been around for a long time, its scheduling capabilities are powerful and proven. They are able to create firing schedules such as: "7:15am every weekday" (0 15 7 ? * MON-FRI) or "10:15am on every last friday of every month during the years 2002 to 2005" (0 15 10 ? * 6L 2002-2005). Cron expressions are powerful, but can be pretty confusing. Fortunately, end users do not need to know about them. The task of defining schedules to the system falls to the subscription administrator. In the near future there will be UI tools to help build the cron expressions.

02. Configuring the Subscription Subsystem

This page last changed on Mar 13, 2007 by gmoran.

The subscription subsystem is part of the additional functionality that previously was included with the Pentaho BI Suite Professional Edition, but is now available in the open source version. It is enabled by default and there is little setup required.

If you are upgrading from a previous Open Source version of the platform and want to modify your existing **pentaho.xml** configuration file. Make sure the following lines are set:

Add the setting 'ISubscriptionRepository' under the 'objects' node:

```
<objects>
<ISubscriptionRepository
scope="global">com.pentaho.repository.subscribe.SubscriptionRepository</ISubscriptionRepository>
.
.
.
</objects>
```

Add the setting 'ProHandler' Under the 'HibernatedObjectHandlers' node:

```
<HibernatedObjectHandlers>
<StandardHandler>org.pentaho.repository.StdHibernateClassHandler</StandardHandler>
<ProHandler>com.pentaho.repository.subscribe.SubscriptionHibernateHandler</ProHandler>
</HibernatedObjectHandlers>
```

Add the setting 'pentaho-pro' Under the 'system-listeners' node:

```
<system-listeners>
<pentaho-pro>com.pentaho.system.ProPentahoSystem</pentaho-pro>
.
.
.
</system-listeners>
```

Subscriptions Basics

Action Sequences are the building blocks of Pentaho BI solutions. They can be executed many different ways. They can be called directly from a URL or via a workflow engine; they can be scheduled or started as part of a web service. They can even be ran as part of a standalone application. The Subscriptions module provides the method and the tools for allowing end users to select, customize and schedule the execution of Action Sequences.

Selection

Action Sequences can do anything from displaying a message; the HelloWorld sample, to bursting out thousands of reports as part of a larger workflow. With such a wide range of functionality, a subscription

doesn't make sense for all types of content. You wouldn't want to subscribe to an Action Sequence that starts a burst job for instance. A more typical subscription would be an Action Sequence that runs a report and emails the result to a user.

In order to manage what content a user subscribes to, a subscription administrator selects which Action Sequences are available for subscription. Each Action Sequence selected may be assigned a set of schedules that the user can select. This selection is performed via the **Subscription Admin** utility.

When a piece of content is defined as being subscribable, a subscribe button will appear on the parameter page when a user executes that Action Sequence. For example, assume the HelloEmail sample Action Sequence in the getting-started group were marked as subscribe-able. When a user clicks on the link to execute HelloEmail, the normal prompt to enter an email address will be presented. In addition, there will be a button labeled "Subscribe" that when clicked shows the subscription options. The user can give it a name, and select a schedule. If they already have subscriptions to this Action Sequence, they can edit, delete or create a new subscription.

Customization

A subscription parameter page allows the user to select the parameter values that will be passed to the subscribed to Action Sequence at runtime. They can set parameters like REGION, Department, email address etc. The parameter page is the same one that is displayed when a user is just running the Action Sequence on its own.

Scheduling

After a user has selected the content to subscribe to and setup the input parameters, they select a schedule for it to run. They are presented with a set of friendly schedule names to choose from. These friendly names are pre-defined by the solution administrator and may be different for each piece of content.

The solution administrator sets up the available schedules by creating a friendly name like "Monday morning" assigns an appropriate cron string to that name and then selects the content that has permission to use that schedule.

Why can't the user create any schedule they want? Allowing that much flexibility introduces a huge amount of uncertainty into the system. A user could, either accidentally or on purpose, run a large report every 5 seconds forever. In most cases there are times that make the most sense to schedule an Action Sequence to run, like after all stores upload their sales figures. In other cases that data may not change for a week or month - so reporting hourly would not make much sense. The solution developer and solution administrator can define as many schedules that make sense for a specific Action Sequence. The users gets to choose from a name that makes sense to them, and the administrator can schedule the run to occur at a time of minimal load.

03. Subscription Admin

This page last changed on Feb 15, 2007 by wgorman.

Overview

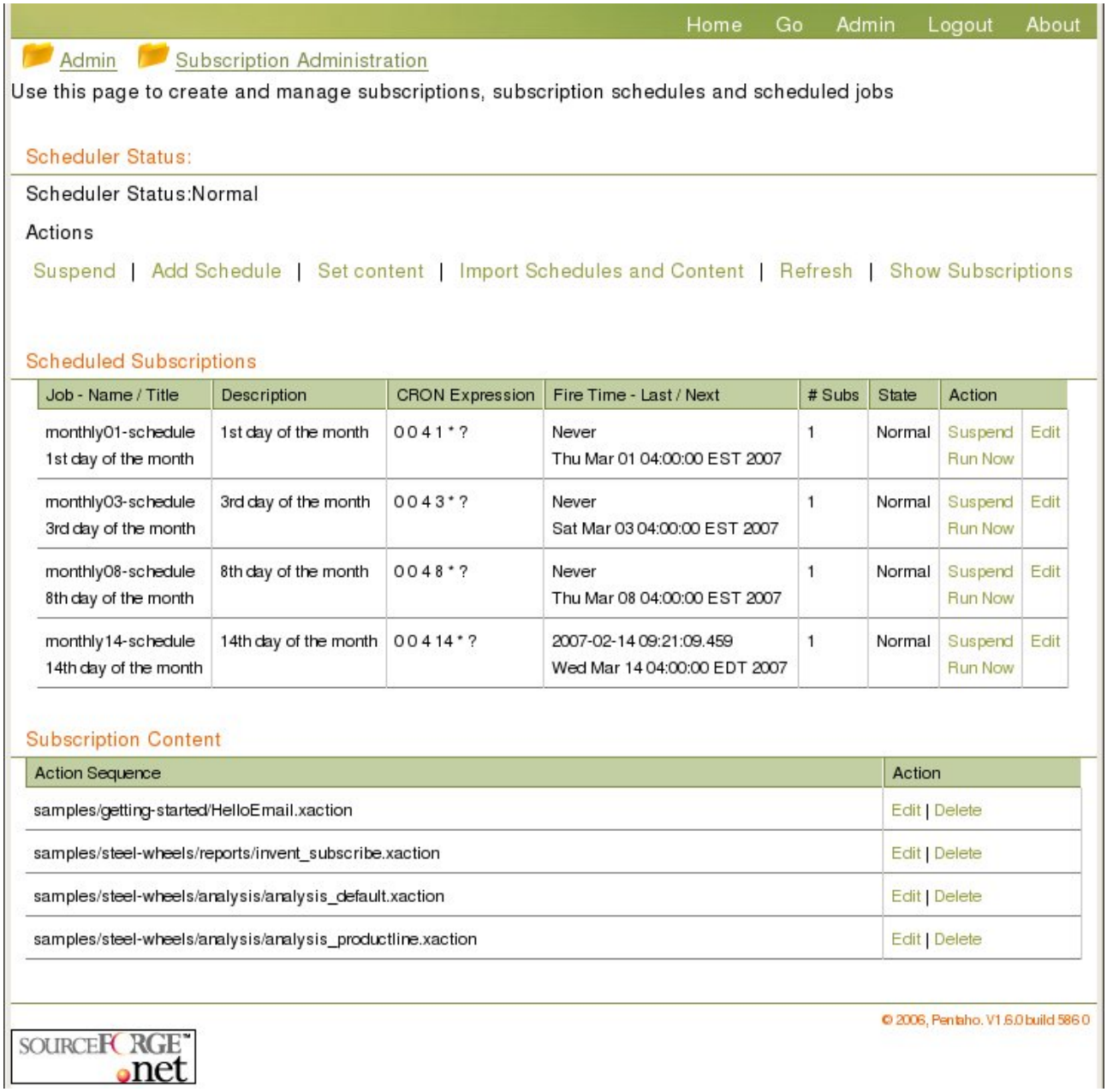
Action Sequences can do anything from displaying a message; the HelloWorld sample, to bursting out thousands of reports as part of a larger workflow. With such a wide range of functionality, a subscription doesn't make sense for all types of content. You wouldn't want to subscribe to an Action Sequence that starts a burst job for instance. A more typical subscription would be an Action Sequence that runs a report and emails the result to a user.

In order to manage what content a user subscribes to, a subscription administrator selects which Action Sequences are available for subscription. Each Action Sequence selected may be assigned a set of schedules that the user can select. This selection is performed via the **Subscription Admin** utility documented here.

When a piece of content is defined as being subscribable, a subscribe button will appear on the parameter page when a user executes that Action Sequence. For example, assume the HelloEmail sample Action Sequence in the getting-started group were marked as subscribe-able. When a user clicks on the link to execute HelloEmail, the normal prompt to enter an email address will be presented. In addition, there will be a button labeled "Subscribe" that when clicked shows the subscription options. The user can give it a name, and select a schedule. If they already have subscriptions to this Action Sequence, they can edit, delete or create a new subscription.

Subscription Administration Page

On the main subscription administration page, a list of Commands, Scheduled Subscriptions, and Subscription Content are displayed.



Available Commands

Suspend / Resume - This command will suspend or resume the entire subscription system from running.

Add Schedule - This command will take the administrator to the Add New Schedule page

Set content - This command will take the administrator to the Set content, where specific action sequences can be added and removed from subscribable content.

Import Schedules and Content - This command will take the administrator to the Subscription Schedule and Content Import page, where XML files can be loaded with pre-configured schedules and content.

Refresh - This command simply reloads the current page.

Show Subscriptions - This command will take the administrator to the Subscriptions page, which lists all existing user subscriptions.

Scheduled Subscriptions

Scheduled Subscriptions are a list of times when subscriptions may be executed. The administrator may select a subset of these schedules for each subscription content. The **# Subs** column represents the number of user subscriptions that will be executed with this schedule. The available commands for Scheduled Subscriptions include:

Suspend / Resume - This command suspends or resumes the execution of any future scheduled user subscriptions using this schedule.

Run Now - This command will execute all user subscriptions mapped to this schedule.

Edit - This command will take the administrator to the Schedule Editor page of this schedule.

Subscription Content

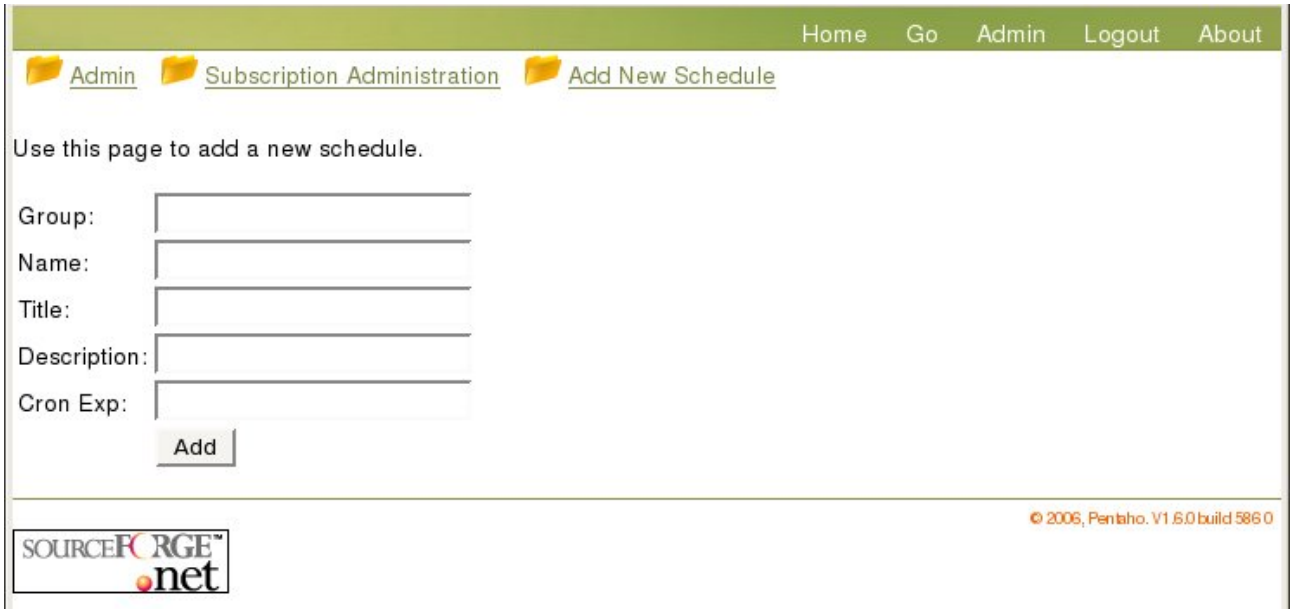
The Subscription Content table lists the action sequences that are currently subscribable. The available commands for Subscription Content include:

Edit - This command will take the administrator to the Subscription Content Editor page for this action sequence.

Delete - This command will remove the action sequence from subscribable content.

Add New Schedule Page

This page allows the administrator to add a new schedule.

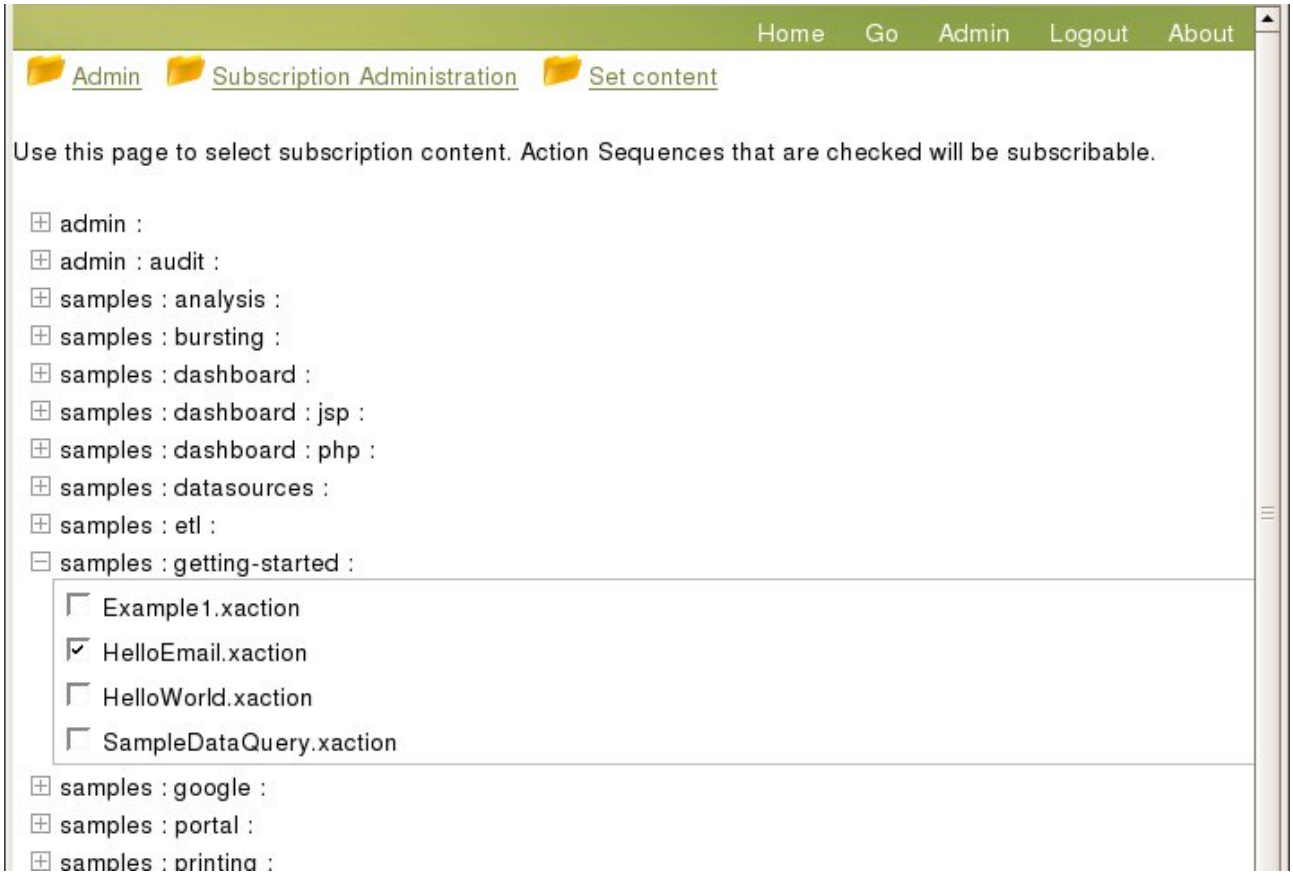


The following fields should be populated:

Group - Not sure how this is used at the moment
Name - The system name for this schedule
Title - The user displayed name for this schedule
Description - The description for the schedule
Cron Exp - A cron expression defining when the schedule should be executed. See the cron expressions page for more information.

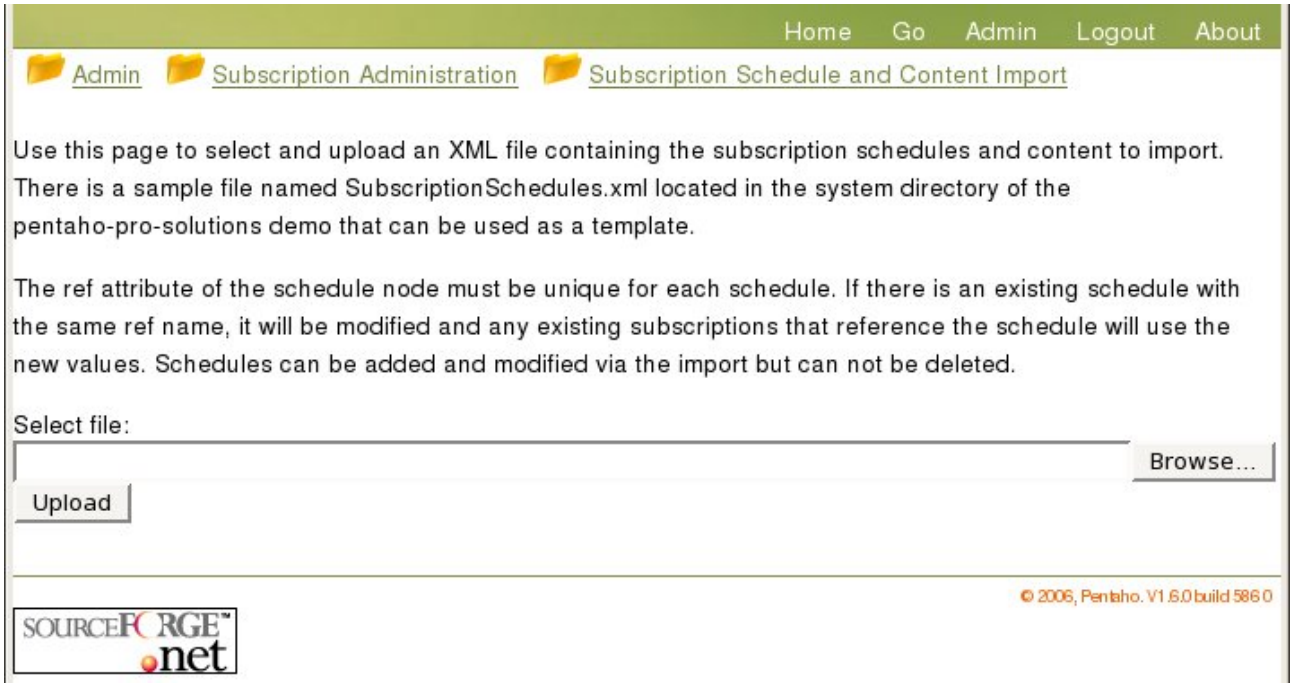
Set Content Page

This page displays a tree of all the available action sequences. The subscription administrator chooses here which action sequences are available for subscriptions. Once selected, click "Set Content" to update the list of available action sequences.



Subscription Schedule and Content Import Page

The import page allows the administrator to import an xml file with pre-existing schedules and content. Select the xml file, and click upload for the changes to take effect.



XML Examples:

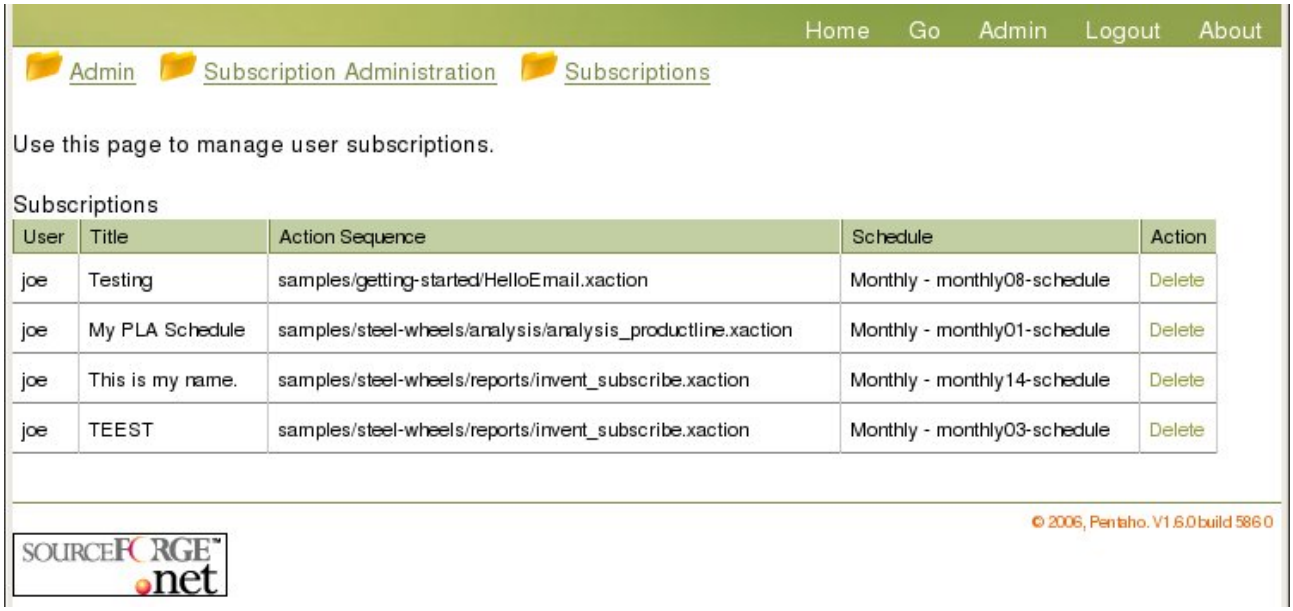
```
<subscription-schedules>
  <schedules>
    <schedule group="Weekly" name="Every Mon at 17:00" description="Every Mon at 5:00 pm"
ref="every1-mon-1700">0 0 17 ? * Mon</schedule>
    <schedule group="Weekly" name="Every Tue at 17:00" description="Every Tue at 5:00 pm"
ref="every2-tue-1700">0 0 17 ? * Tue</schedule>
    <schedule group="Weekly" name="Every Wed at 17:00" description="Every Wed at 5:00 pm"
ref="every3-wed-1700">0 0 17 ? * Wed</schedule>
    <schedule group="Weekly" name="Every Thu at 17:00" description="Every Thu at 5:00 pm"
ref="every4-thu-1700">0 0 17 ? * Thu</schedule>
    <schedule group="Weekly" name="Every Fri at 17:00" description="Every Fri at 5:00 pm"
ref="every5-fri-1700">0 0 17 ? * Fri</schedule>
  </schedules>

  <subscription-content>
    <content action="samples/steel-wheel/HelloEmail.xaction" type="email"
allowAllSchedules="true" />
  </subscription-content>
</subscription-schedules>
```

```
<subscription-schedules>
  <schedules>
   <schedule group="Monthly" name="1st day of the month" description="1st day of the month"
ref="monthly-01-schedule">0 0 4 1 * ?</schedule>
    <schedule group="Monthly" name="Last day of the month" description="Last day of the month"
ref="monthly-last-schedule">0 0 4 L * ?</schedule>
  </schedules>
  <subscription-content>
    <content action="samples/getting-started/HelloEmail.xaction" type="email"
allowAllSchedules="true" />
    <content action="samples/steel-wheels/reports/invent_subscribe.xaction" type="email"
allowAllSchedules="true" />
    <content action="samples/steel-wheels/analysis/analysis_default.xaction" type="email"
allowAllSchedules="true" />
  </subscription-content>
</subscription-schedules>
```

Subscriptions Page

The Subscriptions Page displays a list of all user subscriptions. The administrator may delete these subscriptions from this page.

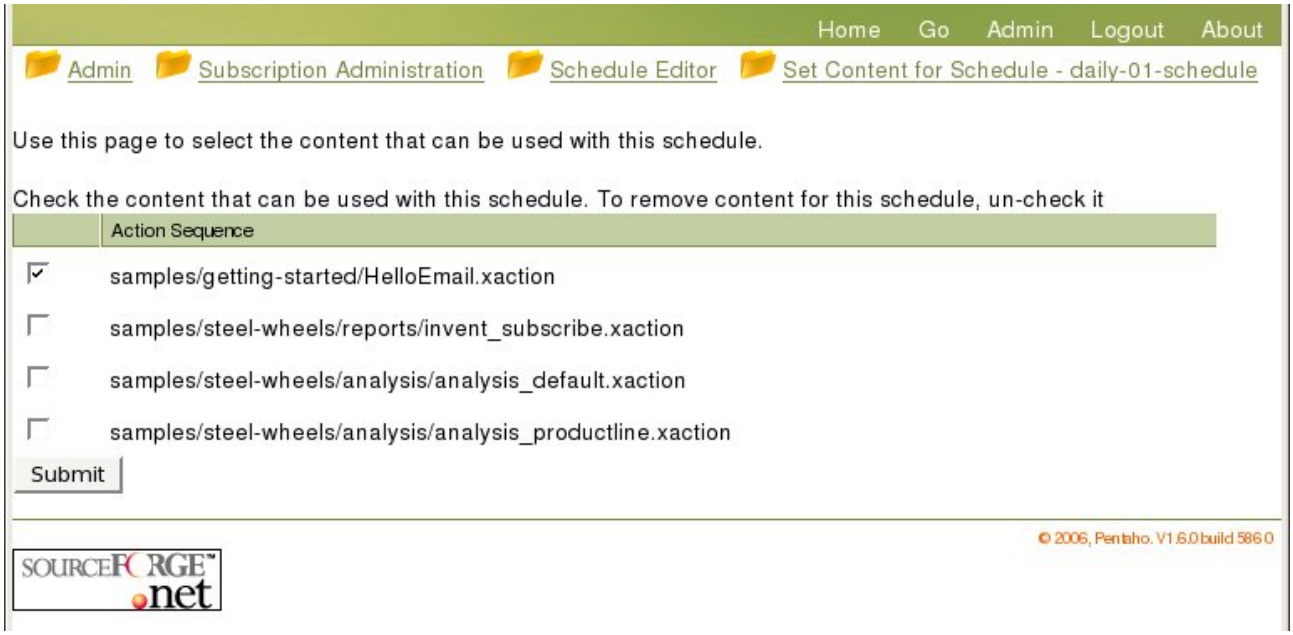


Schedule Editor Page

The schedule editor page allows the administrator to modify the schedule details, and also provides a link to the Set Content for Schedule Page, which allows the administrator to choose which subscription content may be used with this schedule. This page also displays a list of content associated with the schedule.

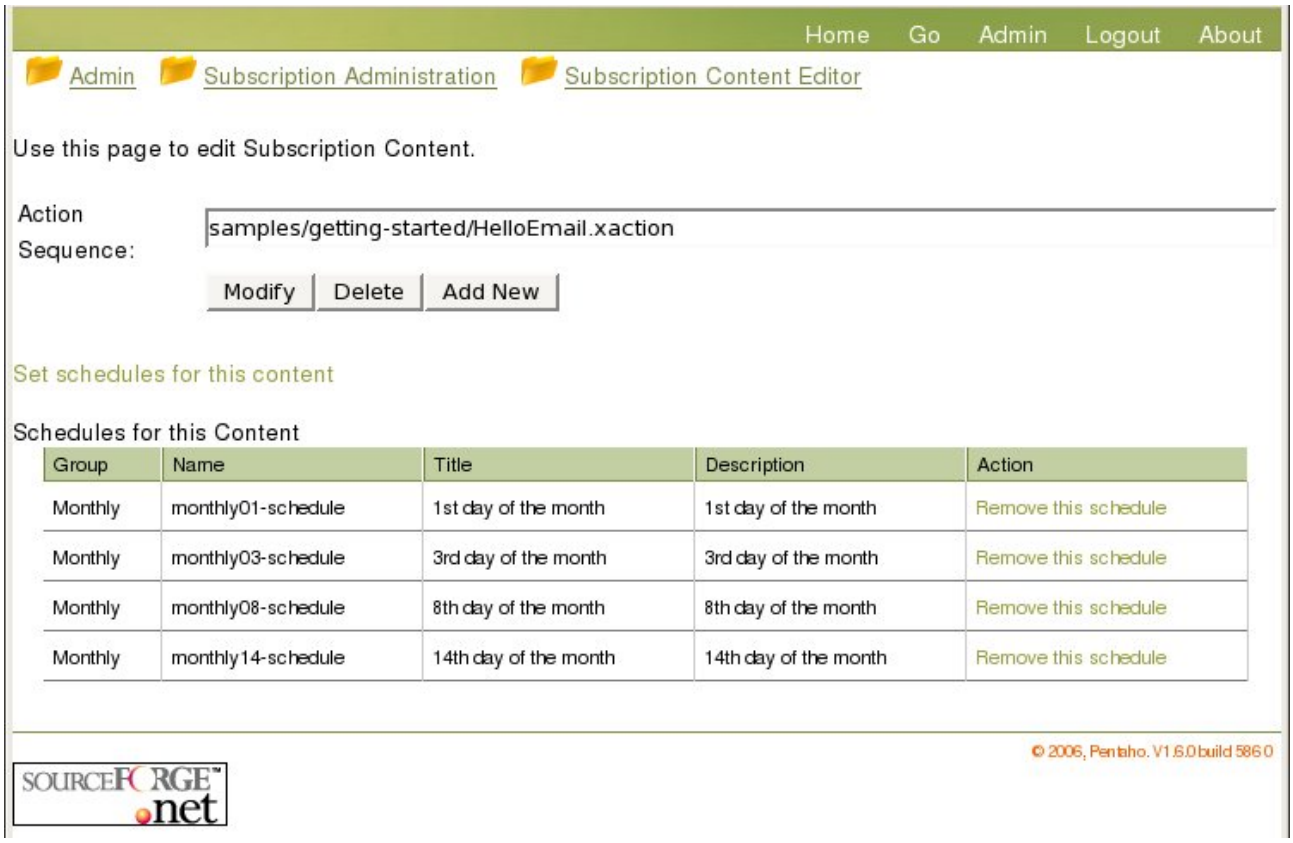
Set Content for Schedule Page

This page allows the administrator to specify which subscription content is associated with a particular schedule.



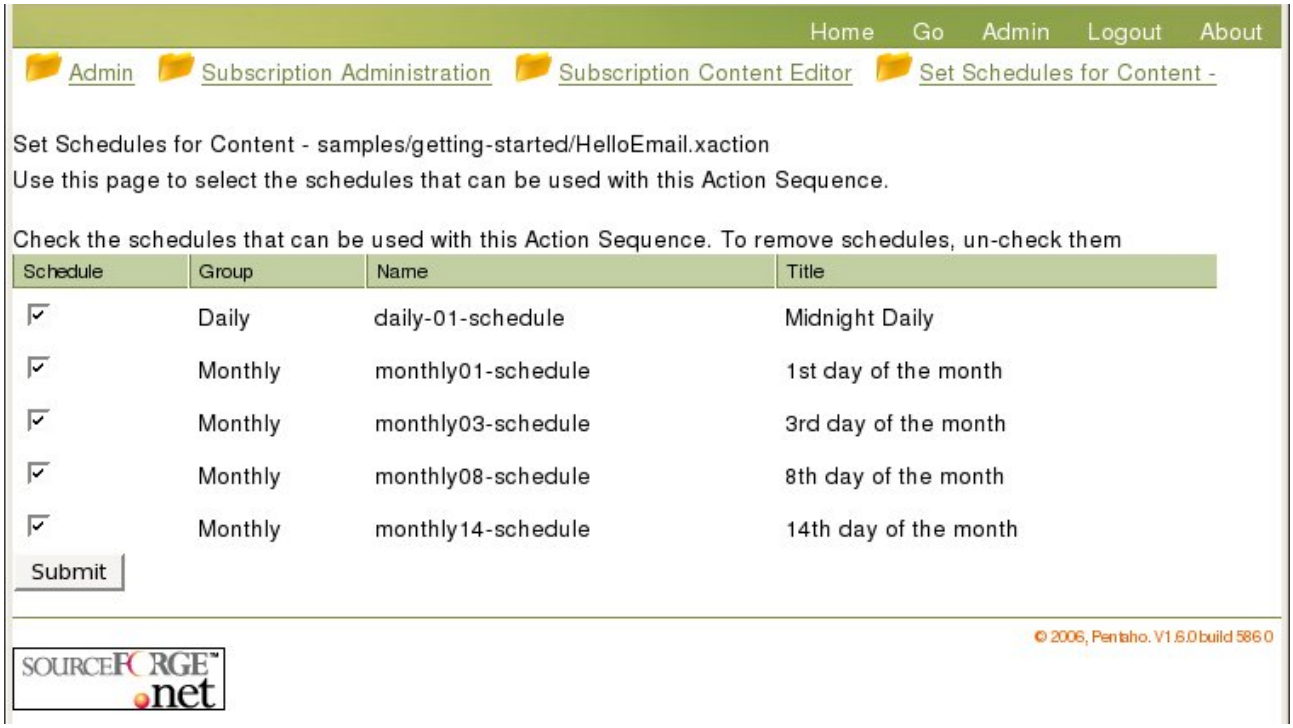
Subscription Content Editor Page

This page allows the administrator to specify a different action sequence for the subscription content. Also this page provides a link to the Set Schedules for Content Page, which allows the administrator to choose which schedules may be used with this content. This page also displays a list of schedules associated with the content.



Set Schedules for Content Page

This page allows the administrator to specify which schedules to associate with a particular subscription content item.



04. User Subscriptions

This page last changed on Feb 15, 2007 by wgorman.

When a piece of content is defined as being subscribable, a subscribe button will appear on the parameter page when a user executes that Action Sequence. For example, assume the HelloEmail sample Action Sequence in the getting-started group was marked as subscribable. When a user clicks on the link to execute HelloEmail, the normal prompt to enter an email address will be presented. In addition, there will be a button labeled "Subscribe" that when clicked shows the subscription options. The user can give it a name, and select a schedule. If they already have subscriptions to this Action Sequence, they can edit, delete or create a new subscription.

Customization

A subscription parameter page allows the user to select the parameter values that will be passed to the subscribed to Action Sequence at runtime. They can set parameters like REGION, Department, email address etc. The parameter page is the same one that is displayed when a user is just running the Action Sequence on its own.

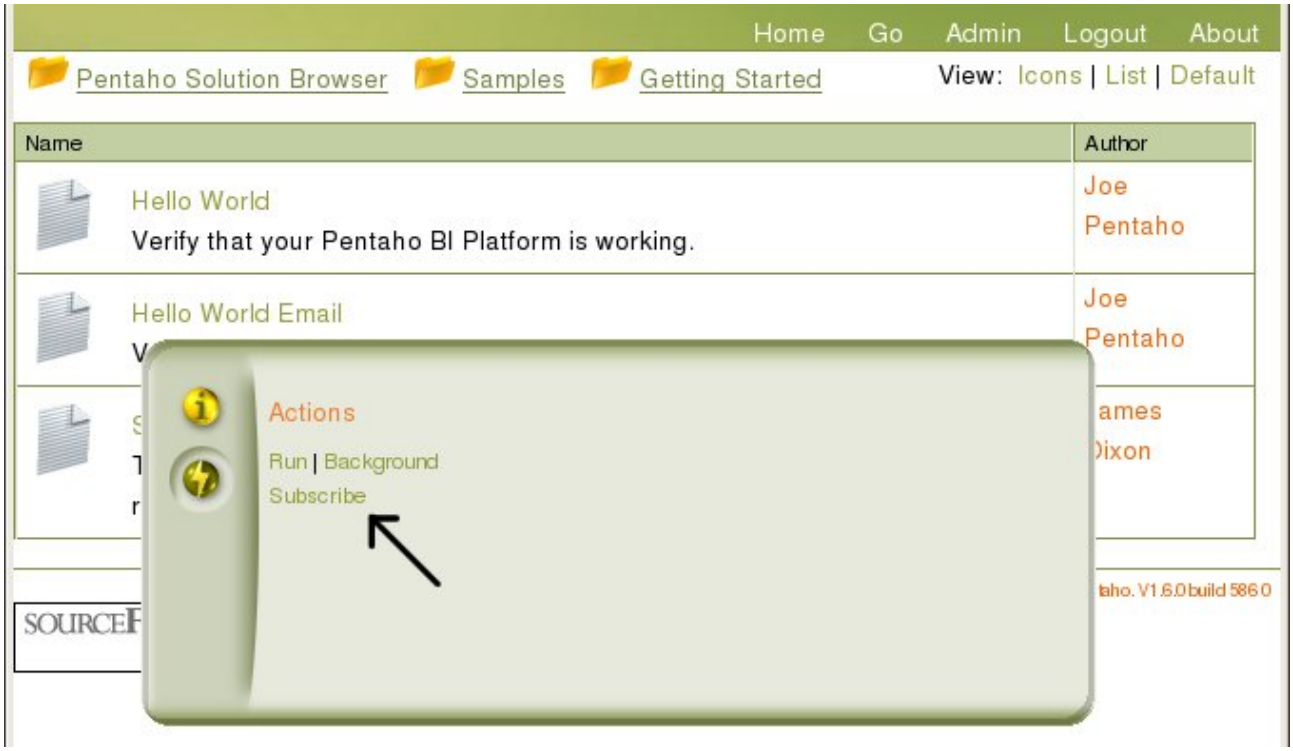
Scheduling

After a user has selected the content to subscribe to and setup the input parameters, they select a schedule for it to run. They are presented with a set of friendly schedule names to choose from. These friendly names are pre-defined by the solution administrator and may be different for each piece of content.

Why can't the user create any schedule they want? Allowing that much flexibility introduces a huge amount of uncertainty into the system. A user could, either accidentally or on purpose, run a large report every 5 seconds forever. In most cases there are times that make the most sense to schedule an Action Sequence to run, like after all stores upload their sales figures. In other cases that data may not change for a week or month - so reporting hourly would not make much sense. The solution developer and solution administrator can define as many schedules that make sense for a specific Action Sequence. The users gets to choose from a name that makes sense to them, and the administrator can schedule the run to occur at a time of minimal load.

Subscribe Button

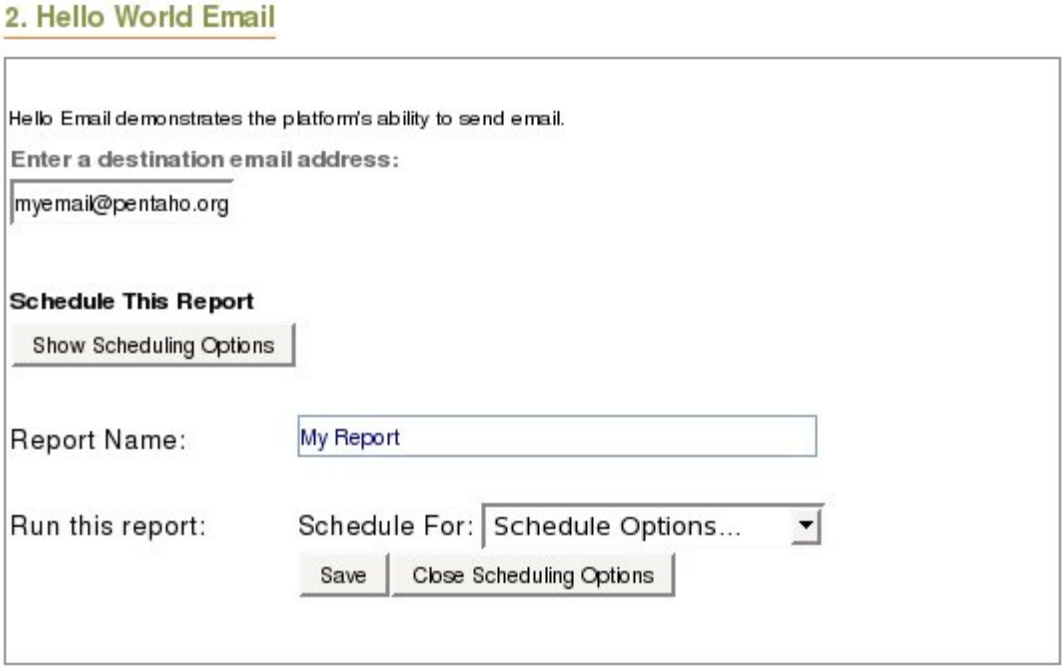
Here is a screenshot of the subscribe button, which appears when an action is subscribable.



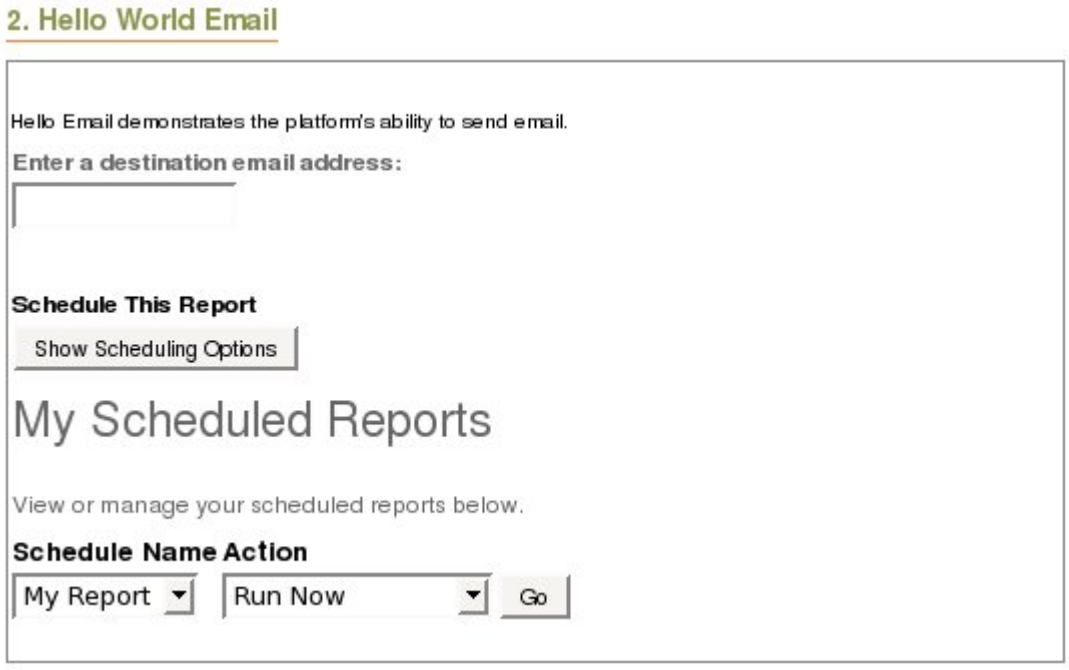
Hello World Email Subscription Example

If the user clicks on the subscribe button, the normal parameter page appears, along with the Show Scheduling Options button.

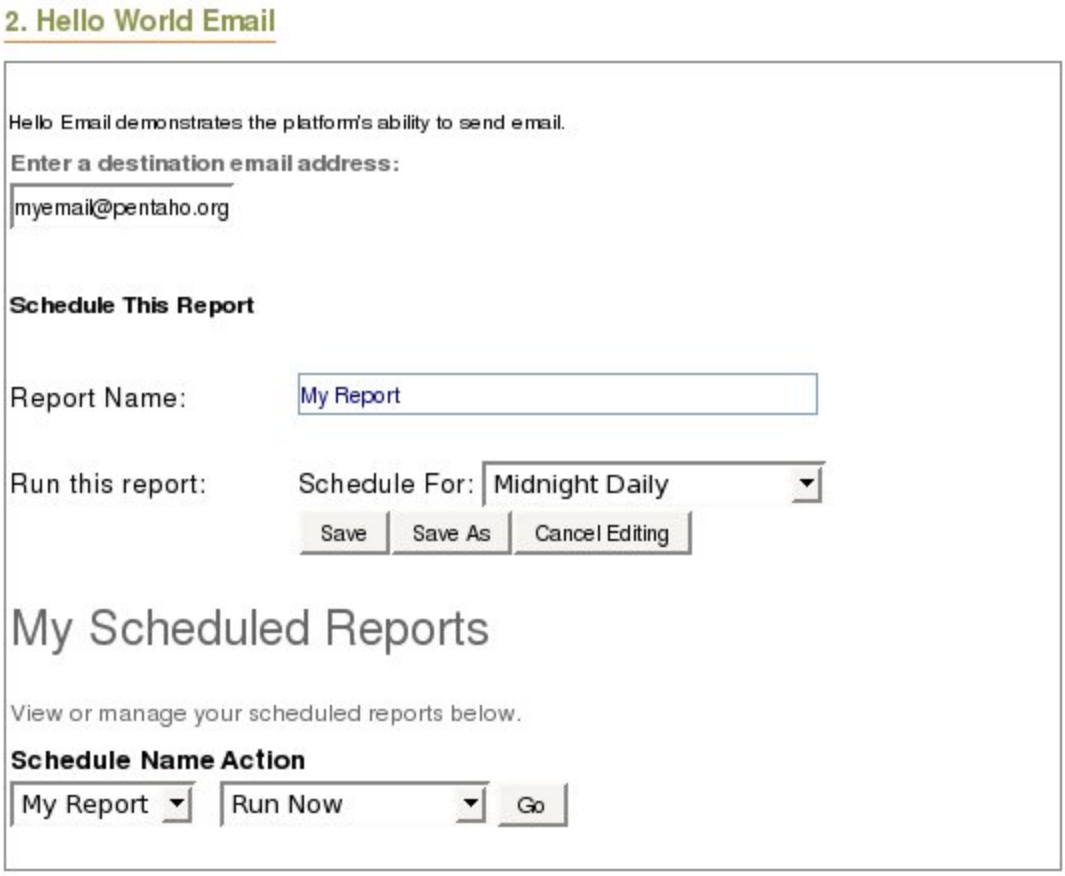
Clicking on the **Show Scheduling Options** button reveals the Report Name field, the **Schedule For** pulldown, and a **Save** button to save the scheduled action.



When returning to this subscription via the above method or via the user subscription page, the user will have the option to run, edit or delete the subscription, and also the ability to add additional subscriptions with different parameters.

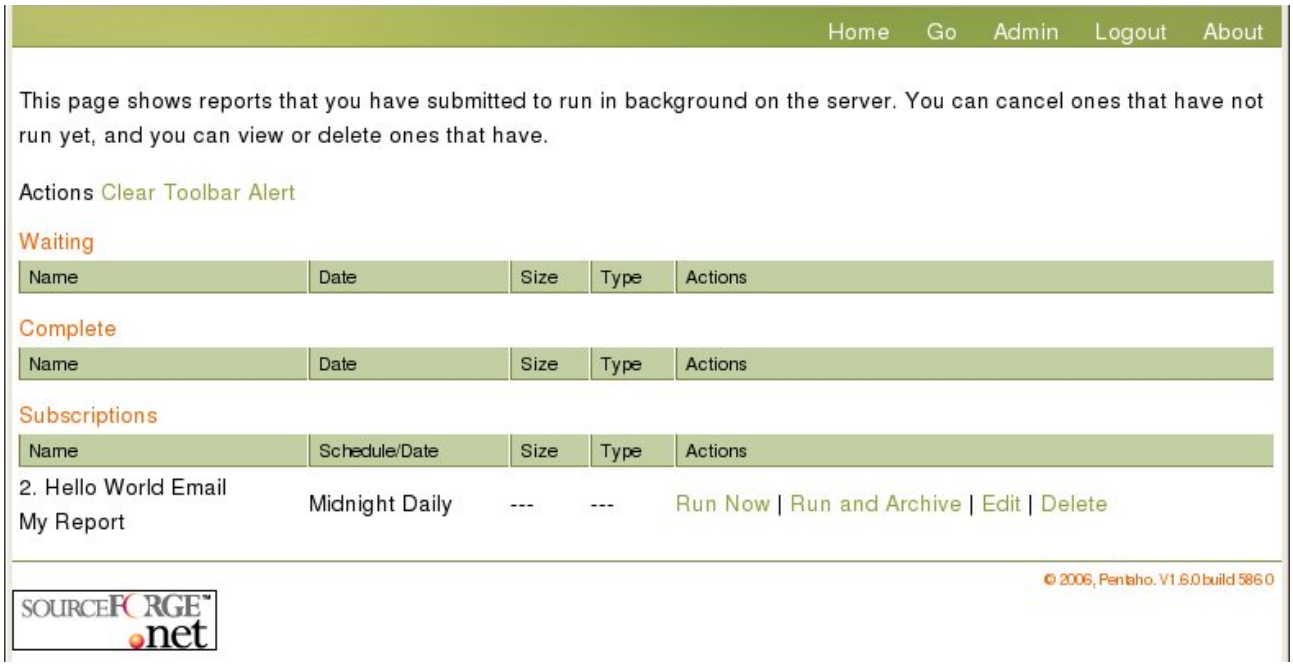


When editing the subscription, the user may change any of the input parameters, save the subscription, and also rename or change the scheduling options.



My Workspace - Subscriptions

From the main menu "Go -> My Workspace" users can view, run, edit, or delete their subscribed actions.



05. CRON Expressions in Detail

This page last changed on Feb 15, 2007 by wgorman.

The following was taken from the CronTriggers Tutorial located on the Quartz website.

Introduction

cron is a UNIX tool that has been around for a long time, so its scheduling capabilities are powerful and proven. The scheduler uses "cron expressions", which are able to create firing schedules such as: "At 8:00am every Monday through Friday" or "At 1:30am every last Friday of the month".

Format

A cron expression is a string comprised of 6 or 7 fields separated by white space. Fields can contain any of the allowed values, along with various combinations of the allowed special characters for that field. The fields are as follows:

Field Name	Mandatory?	Allowed Values	Allowed Special Characters
Seconds	YES	0-59	, - * /
Minutes	YES	0-59	, - * /
Hours	YES	0-23	, - * /
Day of month	YES	1-31	, - * ? / L W C
Month	YES	1-12 or JAN-DEC	, - * /
Day of week	YES	1-7 or SUN-SAT	, - * ? / L C #
Year	NO	empty, 1970-2099	, - * /

So cron expressions can be as simple as this: * * * * ? *
or more complex, like this: 0 0/5 14,18,3-39,52 ? JAN,MAR,SEP MON-FRI 2002-2010

Special characters

- **# ** (*"all values"*) - used to select all values within a field. For example, "*" in the minute field means *"every minute"*.
    1. ?* (*"no specific value"*) - useful when you need to specify something in one of the two fields in which the character is allowed, but not the other. For example, if I want my trigger to fire on a particular day of the month (say, the 10th), but don't care what day of the week that happens to be, I would put "10" in the day-of-month field, and "?" in the day-of-week field. See the examples below for clarification.
    2. -* - used to specify ranges. For example, "10-12" in the hour field means *"the hours 10, 11 and 12"*.

3. ,* - used to specify additional values. For example, "MON,WED,FRI" in the day-of-week field means *"the days Monday, Wednesday, and Friday"*.
4. /* - used to specify increments. For example, "0/15" in the seconds field means *"the seconds 0, 15, 30, and 45"*. And "5/15" in the seconds field means *"the seconds 5, 20, 35, and 50"*. You can also specify '/' after the '**' character - in this case '**' is equivalent to having '0' before the '/'. '1/3' in the day-of-month field means *"fire every 3 days starting on the first day of the month"*.
5. L* (*"last"*) - has different meaning in each of the two fields in which it is allowed. For example, the value "L" in the day-of-month field means *"the last day of the month"*- day 31 for January, day 28 for February on non-leap years. If used in the day-of-week field by itself, it simply means "7" or "SAT". But if used in the day-of-week field after another value, it means *"the last xxx day of the month"* - for example "6L" means *"the last friday of the month"*. When using the 'L' option, it is important not to specify lists, or ranges of values, as you'll get confusing results.
    - **W** (*"weekday"*) - used to specify the weekday (Monday-Friday) nearest the given day. As an example, if you were to specify "15W" as the value for the day-of-month field, the meaning is: *"the nearest weekday to the 15th of the month"*. So if the 15th is a Saturday, the trigger will fire on Friday the 14th. If the 15th is a Sunday, the trigger will fire on Monday the 16th. If the 15th is a Tuesday, then it will fire on Tuesday the 15th. However if you specify "1W" as the value for day-of-month, and the 1st is a Saturday, the trigger will fire on Monday the 3rd, as it will not 'jump' over the boundary of a month's days. The 'W' character can only be specified when the day-of-month is a single day, not a range or list of days.

The 'L' and 'W' characters can also be combined in the day-of-month field to yield 'LW', which translates to *"last weekday of the month"*.

6. #* - used to specify "the nth" XXX day of the month. For example, the value of "6#3" in the day-of-week field means *"the third Friday of the month"*(day 6 = Friday and "#3" = the 3rd one in the month). Other examples: "2#1" = the first Monday of the month and "4#5" = the fifth Wednesday of the month. Note that if you specify "#5" and there is not 5 of the given day-of-week in the month, then no firing will occur that month.
    - **C** (*"calendar"*) - this means values are calculated against the associated calendar, if any. If no calendar is associated, then it is equivalent to having an all-inclusive calendar. A value of "5C" in the day-of-month field means *"the first day included by the calendar on or after the 5th"*. A value of "1C" in the day-of-week field means *"the first day included by the calendar on or after Sunday"*.

The legal characters and the names of months and days of the week are not case sensitive. MON is the same as mon.

Examples

Here are some full examples:

Expression	Meaning
`0 0 12 * * ?`	Fire at 12pm (noon) every day
`0 15 10 ? * *`	Fire at 10:15am every day
`0 15 10 * * ?`	Fire at 10:15am every day
`0 15 10 * * ? *`	Fire at 10:15am every day
`0 15 10 * * ? 2005`	Fire at 10:15am every day during the year 2005

`0 * 14 * * ?`	Fire every minute starting at 2pm and ending at 2:59pm, every day
`0 0/5 14 * * ?`	Fire every 5 minutes starting at 2pm and ending at 2:55pm, every day
`0 0/5 14,18 * * ?`	Fire every 5 minutes starting at 2pm and ending at 2:55pm, AND fire every 5 minutes starting at 6pm and ending at 6:55pm, every day
`0 0-5 14 * * ?`	Fire every minute starting at 2pm and ending at 2:05pm, every day
`0 10,44 14 ? 3 WED`	Fire at 2:10pm and at 2:44pm every Wednesday in the month of March.
`0 15 10 ? * MON-FRI`	Fire at 10:15am every Monday, Tuesday, Wednesday, Thursday and Friday
`0 15 10 15 * ?`	Fire at 10:15am on the 15th day of every month
`0 15 10 L * ?`	Fire at 10:15am on the last day of every month
`0 15 10 ? * 6L`	Fire at 10:15am on the last Friday of every month
`0 15 10 ? * 6L`	Fire at 10:15am on the last Friday of every month
`0 15 10 ? * 6L 2002-2005`	Fire at 10:15am on every last friday of every month during the years 2002, 2003, 2004 and 2005
`0 15 10 ? * 6#3`	Fire at 10:15am on the third Friday of every month
`0 0 12 1/5 * ?`	Fire at 12pm (noon) every 5 days every month, starting on the first day of the month.
`0 11 11 11 11 ?`	Fire every November 11th at 11:11am.

Notes

1. Support for the features described for the 'C' character is not complete.
2. Support for specifying both a day-of-week and a day-of-month value is not complete (you must currently use the '?' character in one of these fields).
3. Be careful when setting fire times between mid-night and 1:00 AM - "daylight savings" can cause a skip or a repeat depending on whether the time moves back or jumps forward.