

Oracle® Fusion Middleware

WebCenter Sites: Analytics Administrator's Guide

11g Release 1 (11.1.1.6.0)

E39242-01

July 2013

Copyright © 2011, 2013 Oracle and/or its affiliates. All rights reserved.

Primary Authors: Melinda Rubenau, Sean Cearley

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate failsafe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

Contents

Preface	vii
Audience	vii
Documentation Accessibility	vii
Related Documents	vii
Conventions	viii
What's New in This Guide	ix
1 Integrating Oracle WebCenter Sites: Analytics with Oracle Web Center Sites	
1.1 About the Integration Process	1-1
1.2 Pre-Integration Checklist	1-2
1.3 Steps for Integrating Analytics with Oracle WebCenter Sites	1-2
1.4 Managing Your Integrated Analytics Installation	1-3
1.5 Granting Users Access to Analytics	1-4
1.6 Managing CM Sites	1-5
1.6.1 Registering CM Sites	1-5
1.6.2 Editing CM Sites	1-6
1.6.3 Deleting CM Sites	1-6
1.7 Managing Asset Reports	1-7
1.7.1 Required Configurations	1-7
1.7.2 Optional Configurations	1-8
1.7.3 Configuring Reports with Custom Parameters	1-8
1.7.4 Configuring an Asset Report	1-8
1.7.5 Deleting an Asset Report	1-10
1.8 Managing Engage Asset Reports	1-11
1.8.1 Managing Recommendations	1-11
1.8.1.1 Registering Recommendations	1-11
1.8.1.2 Editing Recommendations	1-12
1.8.1.3 Deleting Recommendations	1-13
1.8.2 Managing Segments	1-13
1.8.2.1 Registering Segments	1-14
1.8.2.2 Editing Segments	1-14
1.8.2.3 Deleting Segments	1-15
1.9 Managing the Performance Indicator	1-16
1.9.1 Enabling the Performance Indicator	1-16

1.9.2	Disabling the Performance Indicator	1-17
1.10	Managing Users and Groups	1-18
1.10.1	Managing Users	1-18
1.10.1.1	Registering Users	1-18
1.10.1.2	Editing Users	1-19
1.10.1.3	Deleting Users	1-20
1.10.2	Managing User Groups	1-20
1.10.2.1	Creating Groups	1-21
1.10.2.2	Editing Groups	1-21
1.10.2.3	Deleting Groups	1-22
1.10.3	Assigning Users to User Groups	1-22
1.11	Managing Permissions to Reports (and to Analytics)	1-23

2 Configuring Oracle WebCenter Sites: Analytics

2.1	Sensor Overload Alerts	2-1
2.1.1	Setting an 'Overload Alert' Threshold	2-1
2.1.2	Responding to a Red Condition	2-2
2.2	Configuration Files	2-3

3 Hadoop Jobs: Processors and Locations

3.1	Hadoop Jobs Process Flow	3-1
3.2	Monitoring Hadoop Jobs	3-2
3.3	Processors and Locations	3-3
3.4	Object Impressions and Work Packages	3-5
3.4.1	Object Impressions	3-5
3.4.2	Work Packages	3-6
3.4.2.1	Data Collection	3-7
3.4.2.2	Processed Data	3-7
3.4.2.3	Directory Structure for Raw Data and oiprocessed Data	3-7
3.4.2.4	Directory Structure for Daily Work Packages	3-9
3.4.2.5	Directory Structure for Weekly Work Packages	3-9
3.4.2.6	Directory Structure for Monthly Work Packages	3-10
3.5	Processor Descriptions	3-10
3.5.1	Object Impression Processors	3-11
3.5.1.1	OIPProcessor	3-11
3.5.1.1.1	Input Location for OIPProcessor	3-11
3.5.1.1.2	Output Location for OIPProcessor	3-12
3.5.1.2	OIInjection	3-12
3.5.1.2.1	Input Location for OIInjection Processor.	3-13
3.5.1.2.2	Output Locations for OIInjection Processor	3-13
3.5.2	Session Data Processors	3-13
3.5.2.1	SessionMerger	3-13
3.5.2.1.1	Input Location for SessionMerger Processor	3-14
3.5.2.1.2	Output Location for SessionMerger Processor	3-14
3.5.2.2	SessionProcessor	3-14
3.5.2.2.1	Input Location for SessionProcessor.	3-14
3.5.2.2.2	Output Location for SessionProcessor	3-15

3.5.2.3	SessionInjection	3-15
3.5.2.3.1	Input Location for SessionInjection Processor.	3-15
3.5.2.3.2	Output Locations for SessionInjection Processor	3-15
3.5.3	Visitor Data Processors	3-16
3.5.3.1	VisitorMerger	3-16
3.5.3.1.1	Input Location for VisitorMerger Processor.	3-16
3.5.3.1.2	Output Location for VisitorMerger Processor	3-16
3.5.3.2	VisitorMonthProcessor	3-17
3.5.3.2.1	Input Location for VisitorMonthProcessor.	3-17
3.5.3.2.2	Output Location for VisitorMonthProcessor	3-17
3.5.3.3	VisitorMonthInjection	3-17
3.5.3.3.1	Input Location for VisitorMonthInjection Processor.	3-17
3.5.3.3.2	Output Locations for VisitorMonthInjection Processor	3-17
3.5.3.4	VisitorProcessor	3-18
3.5.3.4.1	Input Location for VisitorProcessor.	3-18
3.5.3.4.2	Output Location for VisitorProcessor	3-18
3.5.3.5	VisitorInjection	3-18
3.5.3.5.1	Input Location for VisitorInjection Processor.	3-18
3.5.3.5.2	Output Locations for VisitorInjection Processor	3-18
3.5.3.6	VisitorWeekMerger	3-19
3.5.3.6.1	Input Location for VisitorWeekMerger Processor.	3-19
3.5.3.6.2	Output Location for VisitorWeekMerger Processor	3-19
3.5.3.7	VisitorWeekProcessor	3-19
3.5.3.7.1	Input Location for VisitorWeekProcessor.	3-20
3.5.3.7.2	Output Location for VisitorWeekProcessor	3-20
3.5.3.8	VisitorWeekInjection	3-20
3.5.3.8.1	Input Location for VisitorWeekInjection Processor.	3-20
3.5.3.8.2	Output Locations for VisitorWeekInjection Processor	3-20

Glossary

Preface

The *Oracle Fusion Middleware WebCenter Sites: Analytics Administrator's Guide* guide contains information about integrating a newly installed Analytics system with WebCenter Sites to support the generation of reports with statistics quantifying visitors' browsing patterns on websites powered by Oracle WebCenter Sites. This guide also contains information about configuring operating conditions for the Analytics application and monitoring data processing jobs.

Audience

This document is intended for general administrators of Oracle WebCenter Sites and Oracle WebCenter Sites: Analytics.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>.

Access to Oracle Support

Oracle customers have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> or visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

Related Documents

For more information, see the following documents in the Oracle WebCenter Sites 11g Release 1 (11.1.1.6.0) documentation set:

- *Oracle Fusion Middleware Fusion Middleware WebCenter Sites: Analytics User's Guide*
- *Oracle Fusion Middleware WebCenter Sites: Analytics Developer's Guide*
- *Oracle Fusion Middleware WebCenter Sites: Analytics Tag Reference*
- *Oracle Fusion Middleware WebCenter Sites Installation Guide*
- *Oracle Fusion Middleware WebCenter Sites Administrator's Guide*
- *Oracle Fusion Middleware WebCenter Sites Property Files Reference*

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
<i>italic</i>	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
monospace	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

What's New in This Guide

This preface introduces the new Oracle WebCenter Sites: Analytics Administrator's Guide for Oracle WebCenter Sites version 11g Release 1 (11.1.1.8.0). This is a new guide composed of administrator-specific chapters that came from the *Oracle WebCenter Sites: Installation and Configuration Guide for Analytics* for 11g Release 1 (11.1.1), which is now a part of the consolidated *Oracle Fusion Middleware WebCenter Sites Installation Guide* for 11g Release 1 (11.1.1.8.0).

The following table lists the chapters that compose the new *Oracle WebCenter Sites: Analytics Administrator's Guide* for this release and includes a summary of any changes that were made to these chapters for the current release.

Chapters	Summary of Changes
Chapter 1, "Integrating Oracle WebCenter Sites: Analytics with Oracle Web Center Sites"	This chapter was removed from <i>Oracle WebCenter Sites: Installation and Configuration Guide for Analytics</i> for 11g Release 1 (11.1.1) and added to the current <i>Oracle WebCenter Sites: Analytics Administrator's Guide</i> .
Chapter 2, "Configuring Oracle WebCenter Sites: Analytics"	This chapter was removed from <i>Oracle WebCenter Sites: Installation and Configuration Guide for Analytics</i> for 11g Release 1 (11.1.1) and added to the current <i>Oracle WebCenter Sites: Analytics Administrator's Guide</i> .
Chapter 3, "Hadoop Jobs: Processors and Locations"	<ul style="list-style-type: none">■ This chapter was removed from <i>Oracle WebCenter Sites: Installation and Configuration Guide for Analytics</i> for 11g Release 1 (11.1.1) and added to the current <i>Oracle WebCenter Sites: Analytics Administrator's Guide</i>.■ Added a section about sensor overload alerts. See Section 2.1.
"Glossary"	The glossary terms pertaining to the three chapters listed in this table were taken from the "Glossary" in <i>Oracle WebCenter Sites: Installation and Configuration Guide for Analytics</i> for 11g Release 1 (11.1.1) and placed into the "Glossary" for the <i>Oracle WebCenter Sites: Analytics Administrator's Guide</i> .

Integrating Oracle WebCenter Sites: Analytics with Oracle Web Center Sites

Integrating Analytics with WebCenter Sites is required when the Analytics system is newly installed. Part of a bigger process, integration helps enable users to access the Analytics application and generate reports.

This chapter describes the integration process. It also provides procedures for integrating Analytics with WebCenter Sites.

This chapter contains the following sections:

- [Section 1.1, "About the Integration Process"](#)
- [Section 1.2, "Pre-Integration Checklist"](#)
- [Section 1.3, "Steps for Integrating Analytics with Oracle WebCenter Sites"](#)
- [Section 1.4, "Managing Your Integrated Analytics Installation"](#)
- [Section 1.5, "Granting Users Access to Analytics"](#)
- [Section 1.6, "Managing CM Sites"](#)
- [Section 1.7, "Managing Asset Reports"](#)
- [Section 1.8, "Managing Engage Asset Reports"](#)
- [Section 1.9, "Managing the Performance Indicator"](#)
- [Section 1.10, "Managing Users and Groups"](#)
- [Section 1.11, "Managing Permissions to Reports \(and to Analytics\)"](#)

1.1 About the Integration Process

Integrating Analytics with WebCenter Sites is an administrator's job. Integration is one part of a collaborative process that enables users to generate Analytics reports. The other part involves developers to ensure reports are filled with data, a set of statistics quantifying visitors' browsing patterns on websites powered by WebCenter Sites.

Administrators and developers can complete their parts either concurrently or in any order, as long as they work with the same parameters, assets, and asset types. Steps to be taken by administrators and developers are outlined next:

- During the integration process, you, the administrator registers pages and other asset types with Analytics to enable their recognition and inclusion in reports. **The registration process, alone, does not ensure that reports will be filled with data.**

Reports will remain empty unless developers enable data capture on assets of the types you have registered. Developers may also need to complete other steps, depending on the nature of the reports, as described next.

- To enable data capture, developers tag published pages with the `AnalyticsImgTag` in which they specify the assets and asset types on which data must be captured. If a page is left untagged, no data will be collected for that page (even if the types of assets on that page are already registered with Analytics).

In addition to tagging pages, developers need to consider the nature of the reports. When standard data capture parameters are used, captured data is processed automatically by standard Hadoop jobs. If custom data capture parameters are created, developers must add the custom parameters to the Analytics database, create Hadoop jobs to statistically process the captured data, and code reports to display the statistics.

When the set of conditions outlined in this section are satisfied, data capture begins. Statistics for Analytics reports become available at the end of the data processing cycle, typically 24 hours following the start of data capture.

For more information about enabling data capture and processing, the *Oracle Fusion Middleware WebCenter Sites: Analytics Developer's Guide*. The same guide contains a tutorial on creating reports with custom data capture parameters. To start the integration process, continue to [Section 1.2, "Pre-Integration Checklist."](#)

1.2 Pre-Integration Checklist

Integrating Analytics with WebCenter Sites requires experience with the WebCenter Sites administrative interface. If you need to enable generation of reports on WebCenter Sites: Engage assets, you must have a thorough understanding of recommendations and segments.

Before starting the Analytics-WebCenter Sites integration process, ensure you have the following information readily available:

- Which users must be given access to Analytics.
- Which roles must be assigned to the users.
- Which WebCenter Sites asset types, on which content management sites, must be registered with (identified to) Analytics in order to enable the generation of reports in the "General Information," and "Content Information" report groups.
- If Engage assets will be tracked, which segments and recommendations must be enabled for data capture.
- An asset report remains empty until developers enable data capture (and data processing, if necessary) for the assets to be tracked. Data capture (and processing) can be enabled now, or after the integration process is completed. Information about enabling data capture and processing can be found in the *Oracle Fusion Middleware WebCenter Sites: Analytics Developer's Guide*.

1.3 Steps for Integrating Analytics with Oracle WebCenter Sites

To integrate Analytics with WebCenter Sites and its users

1. Create the `Analytics` role on your WebCenter Sites system and grant it to the appropriate WebCenter Sites users and yourself. For instructions, see [Section 1.5, "Granting Users Access to Analytics."](#)

2. Register content management sites with Analytics. For instructions, see [Section 1.6.1, "Registering CM Sites."](#)
3. Configure the Page Views report. This report configures the `Pageview` object on which reports in the General Information and Content Information groups are based. Reports in these groups will not function until you configure the Page Views report.
 - If you wish to track the searches that visitors perform on your site, you can configure an internal searches report.
 - You can also configure a custom report.
For instructions on configuring reports, see [Section 1.7, "Managing Asset Reports."](#)
4. If you are using Oracle WebCenter Sites: Engage and wish to track recommendations and segments, register their asset types with Analytics. For instructions see the following sections:
 - [Section 1.8.1.1, "Registering Recommendations"](#)
 - [Section 1.8.2.1, "Registering Segments"](#)
5. (Optional) Enable the performance indicator for assets of the types you selected for tracking by Analytics. For instructions, see [Section 1.9, "Managing the Performance Indicator."](#)
6. Register CM site users with Analytics. For instructions, see [Section 1.10.1.1, "Registering Users."](#)
7. Create user groups. For instructions, see [Section 1.10.2.1, "Creating Groups."](#)
8. Grant permissions to user groups to access reports. For instructions, see [Section 1.11, "Managing Permissions to Reports \(and to Analytics\)."](#)
9. Assign users to user groups. For instructions, see [Section 1.10.3, "Assigning Users to User Groups."](#)

Note: At this point, users can generate the reports you have configured, but the reports will remain empty until developers complete their steps, as outlined in [Section 1.1, "About the Integration Process."](#) Statistics for Analytics reports become available at the end of the data processing cycle, typically 24 hours following the start of data capture.

10. Once you have integrated Analytics with your WebCenter Sites system, you can use the procedures in this chapter to continue managing your Analytics installation. For instructions, see [Section 1.4, "Managing Your Integrated Analytics Installation."](#)

1.4 Managing Your Integrated Analytics Installation

Once you have integrated Analytics with your WebCenter Sites system, use procedures in the following sections to manage your Analytics installation:

- [Section 1.5, "Granting Users Access to Analytics"](#)
- [Section 1.6, "Managing CM Sites"](#)
- [Section 1.7, "Managing Asset Reports"](#)

- [Section 1.8, "Managing Engage Asset Reports"](#)
- [Section 1.9, "Managing the Performance Indicator"](#)
- [Section 1.10, "Managing Users and Groups"](#)
- [Section 1.11, "Managing Permissions to Reports \(and to Analytics\)"](#)

1.5 Granting Users Access to Analytics

In this step you will create the `Analytics` role, which grants users access to the Analytics installation, and the permissions to access and generate reports. You will then assign the role to users on your WebCenter Sites system.

To grant users access to Analytics

1. Log in to the WebCenter Sites Admin interface as the administrator ([Figure 1-1](#)). For a new installation, the default administrator user name and password are as follows:
 - User name: `fwadmin`
 - Password: `xceladmin`

Figure 1-1 WebCenter Sites Login Screen



2. Create a role named `Analytics`. For instructions, see the *Oracle Fusion Middleware WebCenter Sites Administrator's Guide*.

Note: If the FirstSite II sample site is installed, it is unnecessary to create the `Analytics` role.

3. Assign the `Analytics` role to the WebCenter Sites users who need to have access to Analytics. For instructions, see the *Oracle Fusion Middleware WebCenter Sites Administrator's Guide*.
4. Assign the `Analytics` role to yourself (the logged in administrator). You will need this role in order to complete the steps necessary to integrate Analytics with your WebCenter Sites system.

1.6 Managing CM Sites

The first step towards enabling visitor activity tracking on your website is registering the underlying content management sites with Analytics. Once you have registered one or more sites, you can change their registration data, or unregister them (that is, delete them from Analytics).

This section contains the following procedures:

- [Section 1.6.1, "Registering CM Sites"](#)
- [Section 1.6.2, "Editing CM Sites"](#)
- [Section 1.6.3, "Deleting CM Sites"](#)

1.6.1 Registering CM Sites

This procedure shows you how to register a CM site with Analytics.

To register a CM site with Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

```
http://<host>:<port>/analyticsadmin/Admin
```

2. In the Sites Management pane on the left, click **Register**.
Analytics displays the Add/Edit Site form ([Figure 1–2](#)).

Figure 1–2 Add/Edit Site

The screenshot shows a web form titled "Add/Edit site". It has the following fields and values:

- Name (*)**: FirstSitel1
- Long Name**: (empty)
- Link (*)**: http://...
- Screen Width (*)**: 1024
- Start Date**: 2009-03-18
- Country**: no_set (dropdown menu)

At the bottom of the form is a "Save" button with a lock icon.

3. Fill in the form as follows:
 - a. **Name** – enter the name of the site, as it appears in the WebCenter Sites Admin interface.
 - b. **Long name** – enter a short, informative description of the site.
 - c. **Link** – enter the URL of the site.
 - d. **Screen Width** – enter the target pixel width for which the site was designed.
 - e. **Start Date** – specify when Analytics should start capturing data on this site.

- f. **Country** – select the country targeted by the site. The selection you make here is used to calculate the **PV/MIO** statistic in the Overview report.
4. Click **Save**.

1.6.2 Editing CM Sites

This procedure shows you how to edit a CM site in Analytics.

To edit a registered CM site

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the Sites Management pane on the left, click **View/Edit**.

Analytics displays a list of currently registered sites (Figure 1–3).

Figure 1–3 Currently Registered Sites

List		
ID	Name	Delete
100040	First Site II	

3. In the list of sites, select the desired site.

Analytics displays an Add/Edit Site form populated with values for the selected site (Figure 1–4).

Figure 1–4 Add/Edit Site

Add/Edit site

Name (*):

Long Name:

Link (*):

Screen Width (*):

Start Date:

Country:

 Save

4. In the form, make your changes, then click **Save**.

1.6.3 Deleting CM Sites

This procedure shows you how to delete a CM site from Analytics.

To delete a CM site from Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) interface via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the Sites Management pane on the left, click **View/Edit**.
Analytics displays a list of currently registered sites (Figure 1–5).

Figure 1–5 Currently Registered Sites

List		
ID	Name	Delete
100040	First Site II	

3. In the list of sites, click the **Delete** (trash can) icon next to the site you want to delete.
4. In the pop-up warning dialog that appears, click **OK**.

1.7 Managing Asset Reports

Before users can generate reports on WebCenter Sites assets, you must first configure the reports. Configuring an asset report enables Analytics to:

- Recognize WebCenter Sites asset types that you register for the report
- Configure report menu options in the Content Information report groups
- Generate reports on assets of the registered types

This section contains the following topics:

- [Section 1.7.1, "Required Configurations"](#)
- [Section 1.7.2, "Optional Configurations"](#)
- [Section 1.7.4, "Configuring an Asset Report"](#)
- [Section 1.7.5, "Deleting an Asset Report"](#)

1.7.1 Required Configurations

After you install Analytics, you must configure the Page Views report. When you configure this report, you configure the `Pageview` object, a system default that specifies the type (or types) of assets Analytics will track. Configuring the `Pageview` object enables default reports that are based on the `Pageview` object.

The following reports can function only when the `Pageview` object is configured:

- Page Views
- Site Information
- Clickstream

For more information about the `Pageview` object, see the glossary at the end of this guide. For instructions on configuring the Page Views report, follow the steps in [Section 1.7.4, "Configuring an Asset Report."](#)

Note: Once you configure the Page Views report, users will be able to generate the reports. To ensure the reports are filled with data, contact developers to enable data capture on assets of the types you specify in the `Pageview` object.

1.7.2 Optional Configurations

If you want Analytics to capture data on the keywords and phrases visitors enter into the site's internal search engine, configure an internal searches report (which will be listed in the Content Information group.) For instructions on configuring reports on internal searches, follow the steps in [Section 1.7.4, "Configuring an Asset Report."](#)

1.7.3 Configuring Reports with Custom Parameters

Reports with custom parameters are configured in the same way as reports with default parameters. Custom parameters will be included in the reports only if developers complete the steps that are outlined in [Section 1.1, "About the Integration Process."](#)

For instructions on configuring a custom asset report, follow the steps in [Section 1.7.4, "Configuring an Asset Report."](#)

1.7.4 Configuring an Asset Report

In this procedure, you will name the asset report, and specify the types of assets that Analytics will track for the report.

Note: Remember the following:

- Reports that you configure will be part of the Content Information report group.
- Configure reports judiciously:

Deleting reports from the system (if necessary) is not trivial. A report deleted from the registration screen is not deleted as a report option. It remains as a menu item in the user's report generation interface and must be removed programmatically.

Deleting a report from the registration screen disables Analytics from analyzing data for that report. Previously analyzed report data is not deleted.

If you need to delete a report from the Administration interface, follow the steps in [Section 1.7.5, "Deleting an Asset Report."](#)

To configure a report (register one or more asset types)

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

```
http://<host>:<port>/analyticsadmin/Admin
```
2. In the Reports on Assets section of the Asset Management pane on the left, click **Configure**.
Analytics displays the Reports on Assets form ([Figure 1–6](#)).

Figure 1–6 Reports on Assets

Report Name	Assets	Options
Page Views	CONTENT, PAGE, PRODUCT	<input type="checkbox"/> isSearch
Content Report	CONTENT_C	<input type="checkbox"/> isSearch
Media Report	MEDIA_C	<input type="checkbox"/> isSearch

Create a new asset report:

isSearch

3. In the form, do the following:
 - a. In the **Site** drop-down list, select the target site.
 - b. In the Report Name column, go to the **Create a new asset report** field, and enter a short, descriptive name for the report.

Note: Remember the following:

- If you are configuring a Page Views report, enter **Page Views**. Analytics will not recognize the report under any other name.
- If you are configuring an internal searches report, enter a name for your report and select the **isSearch** check box (in the Options column).

- c. In the **Assets** field, enter the name(s) of the asset type(s) whose assets will be tracked for this report.

- If you are configuring the Page Views report, use the following convention:

`AssetType1, AssetType2, . . . , AssetTypeN`

Note: Remember the following:

- Enter asset type names exactly as they appear in the WebCenter Sites database.
- Instead of commas, you can also use semicolons (;) to separate asset types in the list.

- If you are configuring an internal searches report, enter *Search* in the **Assets** field.

4. Click **Save**.
5. Next steps:
 - a. (Optional) Enable the performance indicator for assets of the types you selected for tracking by Analytics. For instructions, see [Section 1.9.1, "Enabling the Performance Indicator."](#)
 - b. Grant permissions to users to generate the report you configured:

- Register CM site users with Analytics. For instructions, see [Section 1.10.1.1, "Registering Users."](#)
 - Create user groups. For instructions, see [Section 1.10.2.1, "Creating Groups."](#)
 - Grant permissions to user groups. For instructions, see [Section 1.11, "Managing Permissions to Reports \(and to Analytics\)."](#)
 - Assign users to user groups. For instructions, see [Section 1.10.3, "Assigning Users to User Groups."](#)
- c. An asset report remains empty until data capture is enabled for assets of the specified types. Data capture can be enabled now, or after the integration process is completed. Instructions for enabling data capture can be found in the *Oracle Fusion Middleware WebCenter Sites: Analytics Developer's Guide*.
- d. If developers have created reports with custom parameters, they must also add the parameter to the Analytics database, develop an Analytics job to statistically process data captured on the custom parameter, and code the reports to display the processed data in the desired layout. Instructions are available in the *Oracle Fusion Middleware WebCenter Sites: Analytics Developer's Guide*.

1.7.5 Deleting an Asset Report

This procedure shows you how to delete an asset report from the registration screen (in the Analytics Administration interface).

Note: Remember the following:

- Deleting a report disables Analytics from analyzing data for that report. Previously analyzed report data is not deleted.
 - A report deleted from the registration screen (in the Analytics Administration interface) remains as a menu item in the user's report generation interface and must be removed programmatically.
-
-

To delete an asset report

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Reports on Assets section of the Asset Management pane on the left, click **Configure**.
 Analytics displays the Reports on Assets form ([Figure 1–7](#)).

Figure 1–7 Reports on Assets

Report Name	Assets	Options	
Page Views	CONTENT, PAGE, PRODUCT		
Content Report	CONTENT_C	isSearch	
Media Report	MEDIA_C	isSearch	

Create a new asset report:

isSearch

3. In the form, do the following:
 - a. In the **Site** drop-down list, select the target site.
 - b. Click the **Delete** (trash can) icon next to the report you want to delete.
The report you selected is removed from the Reports on Assets list.

1.8 Managing Engage Asset Reports

This section shows you how to configure Analytics to support Engage assets (recommendations and segments). This allows Analytics to:

- Track usage statistics for recommendations and the assets they reference
- Track segment membership for each visitor

This section contains the following procedures:

- [Section 1.8.1, "Managing Recommendations"](#)
- [Section 1.8.2, "Managing Segments"](#)

1.8.1 Managing Recommendations

This section shows you how to manage recommendations in Analytics. It contains the following procedures:

- [Section 1.8.1.1, "Registering Recommendations"](#)
- [Section 1.8.1.2, "Editing Recommendations"](#)
- [Section 1.8.1.3, "Deleting Recommendations"](#)

1.8.1.1 Registering Recommendations

This procedure shows you how to register recommendations with Analytics. You must register each recommendation you want Analytics to track.

To register a recommendation with Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Recommendations section of the Engage pane on the left, click **Register**.
Analytics displays an empty Add/Edit Recommendation form ([Figure 1–8](#)).

Figure 1–8 Add/Edit Recommendation

3. Fill in the form as follows:

Note: An asterisk (*) next to a field name indicates that the field is required. You cannot leave required fields blank.

- a. In the **Name** field, enter the name of the recommendation asset. The name can, but is not required to, be the same as the name in the WebCenter Sites database.
 - b. In the **Recommendation ID** field, enter the asset ID of the recommendation asset, exactly as it appears in the WebCenter Sites database.
 - c. In the **Site** drop-down list, select the CM site for which you are registering the recommendation.
4. Click **Save**.

1.8.1.2 Editing Recommendations

This procedure shows you how to edit recommendations registered with Analytics.

To edit a recommendation registered with Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Recommendations section of the Engage pane on the left, click **View/Edit**. Analytics displays a list of currently registered recommendations (Figure 1–9).

Figure 1–9 Currently Registered Recommendations

List			
ID	Name	Site	Delete
1124978777055	FSIIHotItems	FirstSiteII	
1125274447712	FSIIHotProducts	FirstSiteII	
0	NoRec	FirstSiteII	

3. In the list, click the desired recommendation.

Analytics displays an Add/Edit Recommendation form populated with values for the selected recommendation (Figure 1–10).

Figure 1–10 Add/Edit Recommendation

4. In the form, make your changes, then click **Save**.

Note: An asterisk (*) next to a field name indicates that the field is required. You cannot leave required fields blank.

1.8.1.3 Deleting Recommendations

This procedure shows you how to delete recommendations from Analytics.

To delete a recommendation from Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
<http://<host>:<port>/analyticsadmin/Admin>
2. In the Recommendations section of the Engage pane on the left, click **View/Edit**.
 Analytics displays a list of currently registered recommendations (Figure 1–11).

Figure 1–11 Currently Registered Recommendations

List			
ID	Name	Site	Delete
1124978777055	FSIIHotItems	FirstSiteII	
1125274447712	FSIIHotProducts	FirstSiteII	
0	NoRec	FirstSiteII	

3. In the list, click the **Delete** (trash can) icon next to the desired recommendation.
4. In the pop-up warning dialog that appears, click **OK**.

1.8.2 Managing Segments

This section shows you how to manage segments in Analytics. It contains the following procedures:

- [Section 1.8.2.1, "Registering Segments"](#)

- [Section 1.8.2.2, "Editing Segments"](#)
- [Section 1.8.2.3, "Deleting Segments"](#)

Note: Before you register segments from your site, you must configure a segment named `NoSeg` with segment ID of 0 (zero), so that visitors that do not belong to any segments can be tracked. You must do this even if all visitors on your site have been assigned to segments.

1.8.2.1 Registering Segments

This procedure shows you how to register segments with Analytics. You must register each segment for which you want to track visitor membership.

To register a segment with Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) interface via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the Segments section of the Engage pane on the left, click **Register**.
Analytics displays an empty Add/Edit Segment form ([Figure 1-12](#)).

Figure 1-12 Add/Edit Segment

3. Fill in the form as follows:

Note: An asterisk (*) next to a field name indicates that the field is required. You cannot leave a required field blank.

- a. In the **Name** field, enter the name of the segment asset. The name can, but is not required to, be the same as in the WebCenter Sites database.
 - b. In the **Segment ID** field, enter the asset ID of the segment asset.
 - c. In the **Site** drop-down list, select the site for which you are registering the segment.
4. Click **Save**.

1.8.2.2 Editing Segments

This procedure shows you how to edit segments registered with Analytics.

To edit a segment registered with Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Segments section of the Engage pane on the left, click **View/Edit**.
 Analytics displays a list of currently registered segments (Figure 1–13).

Figure 1–13 Currently Registered Segments

List			
ID	Name	Site	Delete
1124886357615	AffluentYoungSingles	FirstSiteII	
0	NoSeq	FirstSiteII	

3. In the list, click the desired segment.
 Analytics displays an Add/Edit segment form populated with values for the selected segment (Figure 1–14).

Figure 1–14 Add/Edit Segment

Add/Edit segment

Name (*):

Segment Id (*):

Site (*): ▼

 Save

4. In the form, make your changes, then click **Save**.

Note: An asterisk (*) next to a field name indicates that the field is required. You cannot leave a required field blank.

1.8.2.3 Deleting Segments

This procedure shows you how to delete segments registered with Analytics.

To delete a segment from Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Segments section of the Engage pane on the left, click **View/Edit**.
 Analytics displays a list of currently registered segments (Figure 1–15).

Figure 1–15 Currently Registered Segments

List			
ID	Name	Site	Delete
1124886357615	AffluentYoungSingles	FirstSiteII	
0	NoSeq	FirstSiteII	

- In the list, click the **Delete** (trash can) icon next to the desired segment.
- In the pop-up warning dialog that appears, click **OK**.

1.9 Managing the Performance Indicator

The performance indicator is a bar that shows how often a given asset is visited during a seven-day period, a 30-day period, and a full year. The performance indicator is an option that can be enabled in the Inspect form of every asset being tracked by Analytics. [Figure 1–16](#) shows the performance indicator.

Figure 1–16 Performance Indicator

7d 0	30d 846	365d 846		Peak 514	Trend UP
----------------	-------------------	--------------------	--	--------------------	--------------------

This section shows you how to enable and disable the performance indicator for tagged assets on your WebCenter Sites delivery system. This section contains the following procedures:

- [Section 1.9.1, "Enabling the Performance Indicator"](#)
- [Section 1.9.2, "Disabling the Performance Indicator"](#)

1.9.1 Enabling the Performance Indicator

This procedure shows you how to enable the performance indicator for one or more asset types on your WebCenter Sites system.

To enable the performance indicator

- Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

```
http://<host>:<port>/analyticsadmin/Admin
```

- In the Performance Indicator section of the Asset Management pane on the left, click **Configure**.

Analytics displays the Site form ([Figure 1–17](#)).

Figure 1–17 Site Form

The screenshot shows a web form titled "site". It contains two dropdown menus: "site" with "First Site II" selected, and "piobjecttype" with "CONTENT_C" selected. To the right of the "piobjecttype" dropdown is a checked checkbox labeled "Enable". A "Save" button with a lock icon is located to the right of the "piobjecttype" dropdown.

3. In the form, do the following:
 - a. In the **Site** drop-down list, select the target site.
 - b. In the **piobjecttype** drop-down list, select the desired asset type.
 - c. Select the **Enable** check box.
4. Click **Save**.

1.9.2 Disabling the Performance Indicator

This procedure shows you how to disable the performance indicator for an asset type on your WebCenter Sites system.

To disable the performance indicator

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Performance Indicator section of the Asset Management pane on the left, click **Configure**.

Analytics displays the Site form (Figure 1–18).

Figure 1–18 Site Form

The screenshot shows the same "site" form as Figure 1-17. However, the "Enable" checkbox is now unchecked. The "site" dropdown is still "First Site II" and the "piobjecttype" dropdown is still "CONTENT_C". The "Save" button remains.

3. In the form, do the following:
 - a. In the **Site** drop-down list, select the target site.
 - b. In the **piobjecttype** drop-down list, select the desired asset type.
 - c. Deselect the **Enable** check box.
4. Click **Save**.

1.10 Managing Users and Groups

Access to reports and Analytics interfaces (User and Admin) is granted to WebCenter Sites users through membership in user groups, to which you assign the desired permissions. When integrating Analytics with your WebCenter Sites system, you register the WebCenter Sites users individually, in the Analytics Admin interface, and then add them to a user group.

This section contains the following procedures:

- [Section 1.10.1, "Managing Users"](#)
- [Section 1.10.2, "Managing User Groups"](#)

1.10.1 Managing Users

This section shows you how to manage WebCenter Sites users in Analytics. It contains the following procedures:

- [Section 1.10.1.1, "Registering Users"](#)
- [Section 1.10.1.2, "Editing Users"](#)
- [Section 1.10.1.3, "Deleting Users"](#)

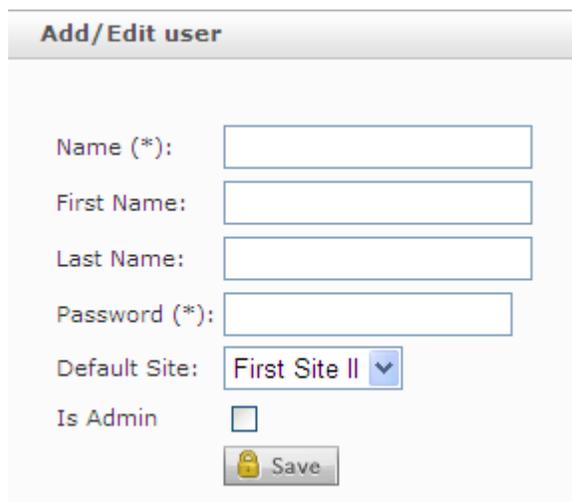
1.10.1.1 Registering Users

This procedure shows you how to register WebCenter Sites users with Analytics.

To register a WebCenter Sites user with Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the User section of the User Management pane on the left, click **Register**.
Analytics displays an empty Add/Edit User form ([Figure 1–19](#)).

Figure 1–19 Add/Edit User Form



The screenshot shows a web form titled "Add/Edit user". It contains the following fields and controls:

- Name (*):** A text input field.
- First Name:** A text input field.
- Last Name:** A text input field.
- Password (*):** A text input field.
- Default Site:** A dropdown menu with "First Site II" selected.
- Is Admin:** An unchecked checkbox.
- Save:** A button with a lock icon and the text "Save".

3. Fill in the form as follows:

- a. **Name (required)** – enter the user name of the user as it appears in the WebCenter Sites interface.
 - b. **First name** – enter the user's first name.
 - c. **Last name** – enter the user's last name.
 - d. **Password** – enter the password you want to assign to the user.
 - e. **Default Site** – select the which the user will see when he/she logs in to the Analytics interface.
 - f. **Is Admin** – grants the user access to the Analytics administration interface.
4. Click **Save**.

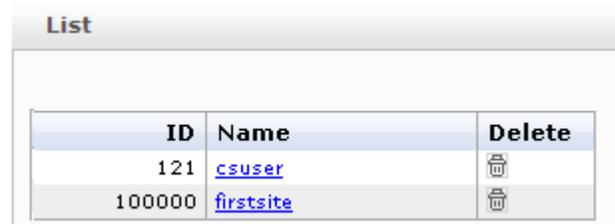
1.10.1.2 Editing Users

This procedure shows you how to edit WebCenter Sites users that have been registered with Analytics.

To edit a registered WebCenter Sites user

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the User section of the User Management pane on the left, click **View/Edit**.
 Analytics displays a list of currently registered users (Figure 1–20).

Figure 1–20 *Currently Registered Users*



List		
ID	Name	Delete
121	csuser	
100000	firstsite	

3. In the list, click the desired user.
 Analytics displays an Add/Edit User form populated with values for the selected user (Figure 1–21).

Figure 1–21 Add/Edit User

Add/Edit user

Name (*):

First Name:

Last Name:

Password (*):

Default Site: ▼

Is Admin

Save

4. In the form, make your changes, then click **Save**.

1.10.1.3 Deleting Users

This procedure shows you how to delete WebCenter Sites users from Analytics.

To delete a WebCenter Sites user from Analytics

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the User section of the User Management pane on the left, click **View/Edit**. Analytics displays a list of currently registered users ([Figure 1–22](#)).

Figure 1–22 Currently Registered Users

ID	Name	Delete
121	csuser	
100000	firstsite	

3. In the list, click the **Delete** (trash can) icon next to the desired user.
4. In the pop-up warning dialog that appears, click **OK**.

1.10.2 Managing User Groups

The purpose of user groups is to grant permissions to group members. This section shows you how to manage user groups in Analytics. It contains the following sections:

- [Section 1.10.2.1, "Creating Groups"](#)
- [Section 1.10.2.2, "Editing Groups"](#)
- [Section 1.10.2.3, "Deleting Groups"](#)

1.10.2.1 Creating Groups

This procedure shows you how to add a user group to Analytics.

To add a user group

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Usergroups section of the User Management pane on the left, click **Register**. Analytics displays an empty Add/Edit Usergroup form (Figure 1–23).

Figure 1–23 Add/Edit Usergroup

3. In the form, enter a name for the group and click **Save**.

1.10.2.2 Editing Groups

This procedure shows you how to edit a user group in Analytics.

To edit a user group

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:
`http://<host>:<port>/analyticsadmin/Admin`
2. In the Usergroups section of the User Management pane on the left, click **View/Edit**. Analytics displays a list of existing user groups (Figure 1–24).

Figure 1–24 Existing User Groups

ID	Name	Delete
10	csgroup	

3. In the list, click the desired user group. Analytics displays the name of the user group you wish to edit in the Add/Edit Usergroup form (Figure 1–25).

Figure 1–25 Add/Edit User Group

4. In the form, make your changes and click **Save**.

1.10.2.3 Deleting Groups

This procedure shows you how to delete a user group from Analytics.

To delete a user group

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the Usergroups section of the User Management pane on the left, click **View/Edit**.

Analytics displays a list of existing user groups (Figure 1–26).

Figure 1–26 Existing User Groups

ID	Name	Delete
10	csgroup	

3. In the list, click the **Delete** (trash can) next to the desired user group.
4. In the pop-up warning dialog that appears, click **OK**.

1.10.3 Assigning Users to User Groups

This procedure shows you how to assign users to user groups in Analytics. The users can access only the reports to which their group has permissions.

To assign users to a user group

1. Log in to the Analytics administration interface as the `csuser` user (password: `csuser`) via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the User to Usergroup section of the User Management pane, click **Register**.

Analytics displays the Usergroup form (Figure 1–27).

Figure 1–27 Usergroup

3. In the form, do the following:
 - a. In the **usergroup** drop-down list, select the target group.
 - b. Select the users that you want to add to the user group. Deselect the users you want to remove from the group.
4. Click **Save**.

1.11 Managing Permissions to Reports (and to Analytics)

When you assign reports to a user group, users in the group gain access to the reports. This procedure shows you how to grant permissions to a user group to access Analytics reports.

Note: At the very least, a user group – and therefore its users must be given access to at least one report on one of the content management sites they work with. Otherwise, they will not be able to log in to the Analytics interface.

To grant or deny users permissions to reports

1. Log in to the Analytics administration interface as the **csuser** user (password: **csuser**) via the following URL:

`http://<host>:<port>/analyticsadmin/Admin`

2. In the Permissions pane on the left, click **Grant Rights**.
Analytics displays the Grant Rights form ([Figure 1–28](#)).

Figure 1–28 Grant Rights Form

The screenshot shows a web form titled "usergroup". It contains two dropdown menus: "usergroup" (set to "csgroup") and "site" (set to "First Site II"). Below these are two columns of checkboxes for selecting reports. The first column lists reports like CONTENT_C, CSREPORT, CS_Assets, CS_Users, Page Views, REPORT, Search, asset, browser, clickstream, country, entryexit, externalsearch, hostname, ip, object, os, and rec. The second column lists reports like CONTENT_C, CSREPORT, and rec. Checkmarks are visible next to CONTENT_C, CSREPORT, hostname, and rec.

3. In the form, do the following:
 - a. In the **Usergroup** drop-down list, select the desired group.
 - b. In the **Site** drop-down list, select the target site.
 - c. Select the reports that the user group may access. Users in the group will be able to access only those reports.
4. Click **Save**.

Configuring Oracle WebCenter Sites: Analytics

Analytics operating conditions can be configured in several files. This chapter discusses configuring the sensor overload threshold in particular, and lists various configuration files where other types of parameters can be set.

This chapter contains the following sections:

- [Section 2.1, "Sensor Overload Alerts"](#)
- [Section 2.2, "Configuration Files"](#)

2.1 Sensor Overload Alerts

During heavy site traffic, the Analytics Sensor can become overloaded with incoming data and stop responding normally. The Analytics Sensor will stop writing to the file system and will instead store incoming data in memory, until an out-of-memory condition is reached.

The Analytics Administrator interface ([Figure 2-1](#)) alerts you to an overload condition by displaying the **Analytics Sensor** button in either yellow or red. Yellow indicates a WARNING condition. Red indicates a CRITICAL condition (assuming the sensor is running).

Note: A stopped or non-functional Analytics Sensor is also displayed in red.

This section contains the following topics:

- [Section 2.1.1, "Setting an 'Overload Alert' Threshold"](#)
- [Section 2.1.2, "Responding to a Red Condition"](#)

2.1.1 Setting an 'Overload Alert' Threshold

Properties in `global.xml` determine the threshold that triggers a WARNING or CRITICAL condition. The properties are `sensor.requestqueue.warnsize` and `sensor.requestqueue.maxsize`. Set these properties to a threshold that is compatible with the configuration of your Analytics installation and the volume of site traffic. For more information about these properties, see the *Oracle Fusion Middleware WebCenter Sites Installation Guide*.

2.1.2 Responding to a Red Condition

If you are monitoring the Status Summary panel (in the Administration interface, [Figure 2-1](#)) and you notice that the **Analytics Sensor** button is displayed in red, you need to determine whether the Sensor has stopped, has failed, or is overloaded. In case of overload, you will need to clear the memory in order to reset the system and resume normal functioning.

Note: Data cleared from memory cannot be retrieved and will be lost.

Figure 2-1 Analytics Administrator Interface, Status Summary Panel

Sensor List		Sensor Details: Sensor	
Identifier	Sensor-saraswatlinux-36		
Status:			
Status indicator	RUNNING		
Status	Processed 0 Requests (Writer #0: 0) in 0ms (avg: 0.0 /sec)		
Exception			
Queue size	0		
Average processing	0.0		
Total OI written	2		
Last OI written on	2009-06-23 16:01:02		
Hardware details:			
Operating System	Linux 2.6.24-23-generic / amd64		
CPUs	2		
Memory (Used / Available) [byte]	137953280 byte / 637136960 byte		
Stop this sensor			
Clear retention pond			

To respond to a red condition

Click the **Analytics Sensor** button and note the main panel.

- If you see "No data available", the Analytics Sensor has either stopped or failed.
- If you see the Sensor Details panel, the Analytics Sensor is running, but it is overloaded. Click the icon labeled **Clear retention pond** to clear the memory.

Figure 2–2 Retention Pond

Components	Sensor List	Sensor Details: Sensor																														
<ul style="list-style-type: none"> Overview Sensors Hdfs Agents Processors Tasks Data Locations Database 		<table border="1"> <tr> <td>Identifier</td> <td>Sensor-saraswatlinux-36</td> </tr> <tr> <td>Status:</td> <td></td> </tr> <tr> <td> Status indicator</td> <td>RUNNING</td> </tr> <tr> <td> Status</td> <td>Processed 0 Requests (Writer #0: 0) in 0ms (avg: 0.0 /sec)</td> </tr> <tr> <td> Exception</td> <td></td> </tr> <tr> <td>Queue size</td> <td>0</td> </tr> <tr> <td>Average processing</td> <td>0.0</td> </tr> <tr> <td>Total OI written</td> <td>2</td> </tr> <tr> <td>Last OI written on</td> <td>2009-06-23 16:01:02</td> </tr> <tr> <td>Hardware details:</td> <td></td> </tr> <tr> <td> Operating System</td> <td>Linux 2.6.24-23-generic / amd64</td> </tr> <tr> <td> CPUs</td> <td>2</td> </tr> <tr> <td> Memory (Used / Available) [byte]</td> <td>137953280 byte / 637136960 byte</td> </tr> <tr> <td>Stop this sensor</td> <td></td> </tr> <tr> <td>Clear retention pond</td> <td></td> </tr> </table>	Identifier	Sensor-saraswatlinux-36	Status:		Status indicator	RUNNING	Status	Processed 0 Requests (Writer #0: 0) in 0ms (avg: 0.0 /sec)	Exception		Queue size	0	Average processing	0.0	Total OI written	2	Last OI written on	2009-06-23 16:01:02	Hardware details:		Operating System	Linux 2.6.24-23-generic / amd64	CPUs	2	Memory (Used / Available) [byte]	137953280 byte / 637136960 byte	Stop this sensor		Clear retention pond	
Identifier	Sensor-saraswatlinux-36																															
Status:																																
Status indicator	RUNNING																															
Status	Processed 0 Requests (Writer #0: 0) in 0ms (avg: 0.0 /sec)																															
Exception																																
Queue size	0																															
Average processing	0.0																															
Total OI written	2																															
Last OI written on	2009-06-23 16:01:02																															
Hardware details:																																
Operating System	Linux 2.6.24-23-generic / amd64																															
CPUs	2																															
Memory (Used / Available) [byte]	137953280 byte / 637136960 byte																															
Stop this sensor																																
Clear retention pond																																
<ul style="list-style-type: none"> Components Sites Management Asset Management Engage User Management Permissions 																																

2.2 Configuration Files

Analytics operating conditions are defined in the following files:

- `global.xml`, an Analytics file that defines installation directories, the handling of raw data, data processing conditions, and system administrators' contact information.
- `log4j.properties`, which defines logging for the Hadoop Job scheduler
- `futuretense_xcel.ini`, a WebCenter Sites property file that contains Analytics properties defining, for example, URLs of the Analytics servlet and generated reports.

For more information about the configuration files, see the *Oracle Fusion Middleware WebCenter Sites Installation Guide*.

Hadoop Jobs: Processors and Locations

Hadoop Jobs consists of locations that store site visitor data and processors that selectively handle the portion of the data they are programmed to handle. A given processor reads one location, processes the data, and writes the results to the next location for pickup by the next processor.

This chapter contains a summary of Hadoop Jobs, guidelines for monitoring Hadoop Jobs, and descriptions of the processors and locations. This chapter contains the following sections:

- [Section 3.1, "Hadoop Jobs Process Flow"](#)
- [Section 3.2, "Monitoring Hadoop Jobs"](#)
- [Section 3.3, "Processors and Locations"](#)
- [Section 3.4, "Object Impressions and Work Packages"](#)
- [Section 3.5, "Processor Descriptions"](#)

3.1 Hadoop Jobs Process Flow

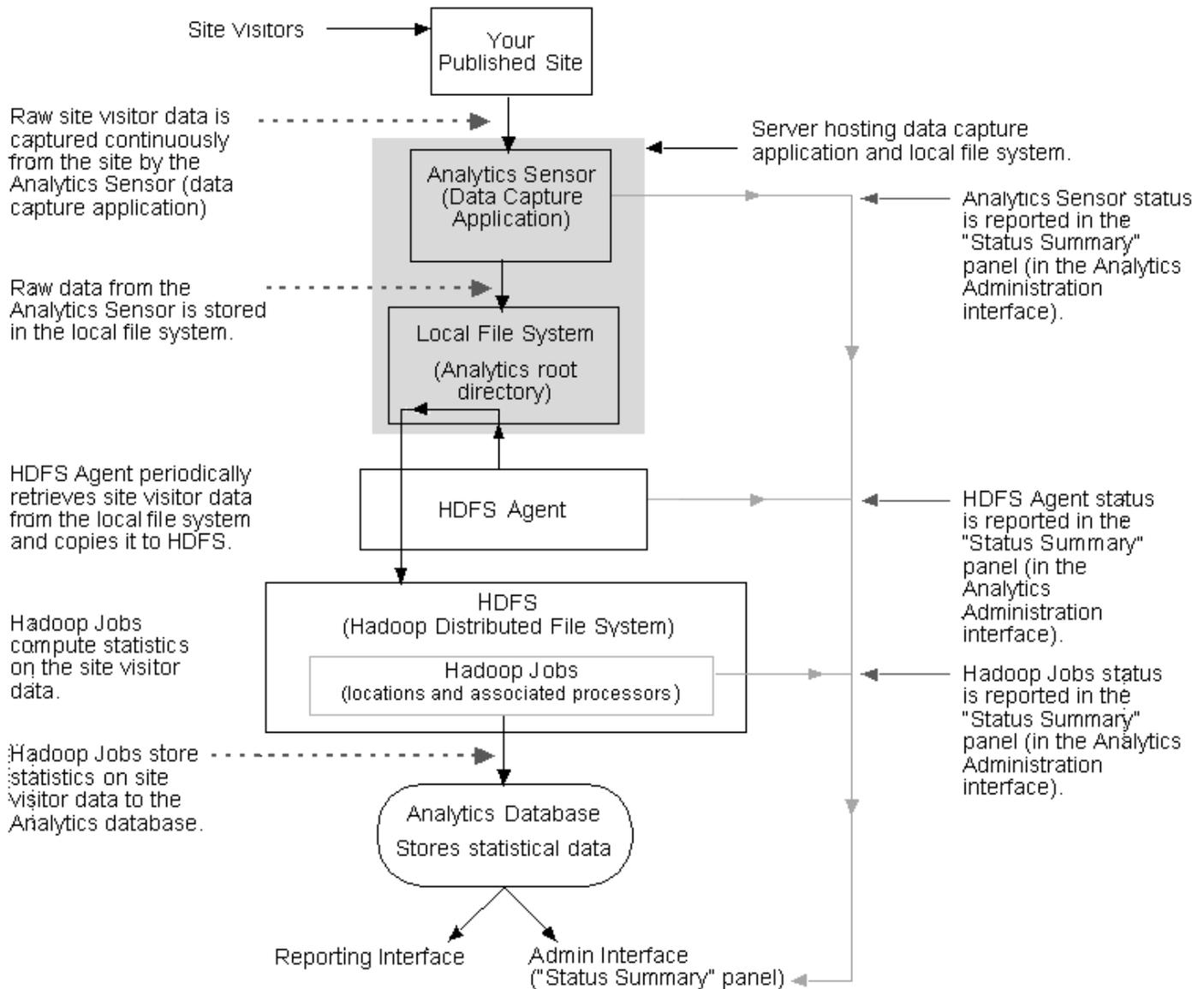
Hadoop Jobs is an Oracle application that statistically processes Analytics data and stores the results in the Analytics database.

In a functional Analytics installation, raw site visitor data is continuously captured by the Analytics Sensor (Data Capture Application), which then stores the data to the local file system. The raw data in the file system is called on periodically by the Hadoop Distributed File System (HDFS) Agent, which then copies the raw data to the Hadoop Distributed File System, where Hadoop jobs process the data.

Hadoop jobs consist of locations and Oracle-specific processors that read site visitor data in one location, statistically process that data, and write the results to another location for pickup by the next processor. When processing is complete, the results (statistics on the raw data) are injected into the Analytics database.

Hadoop jobs can be monitored from the Status Summary panel of the Analytics Administration interface (see [Section 3.2, "Monitoring Hadoop Jobs"](#)). [Figure 3-1](#) depicts Hadoop Jobs process flow.

Figure 3–1 Hadoop Jobs Process Flow



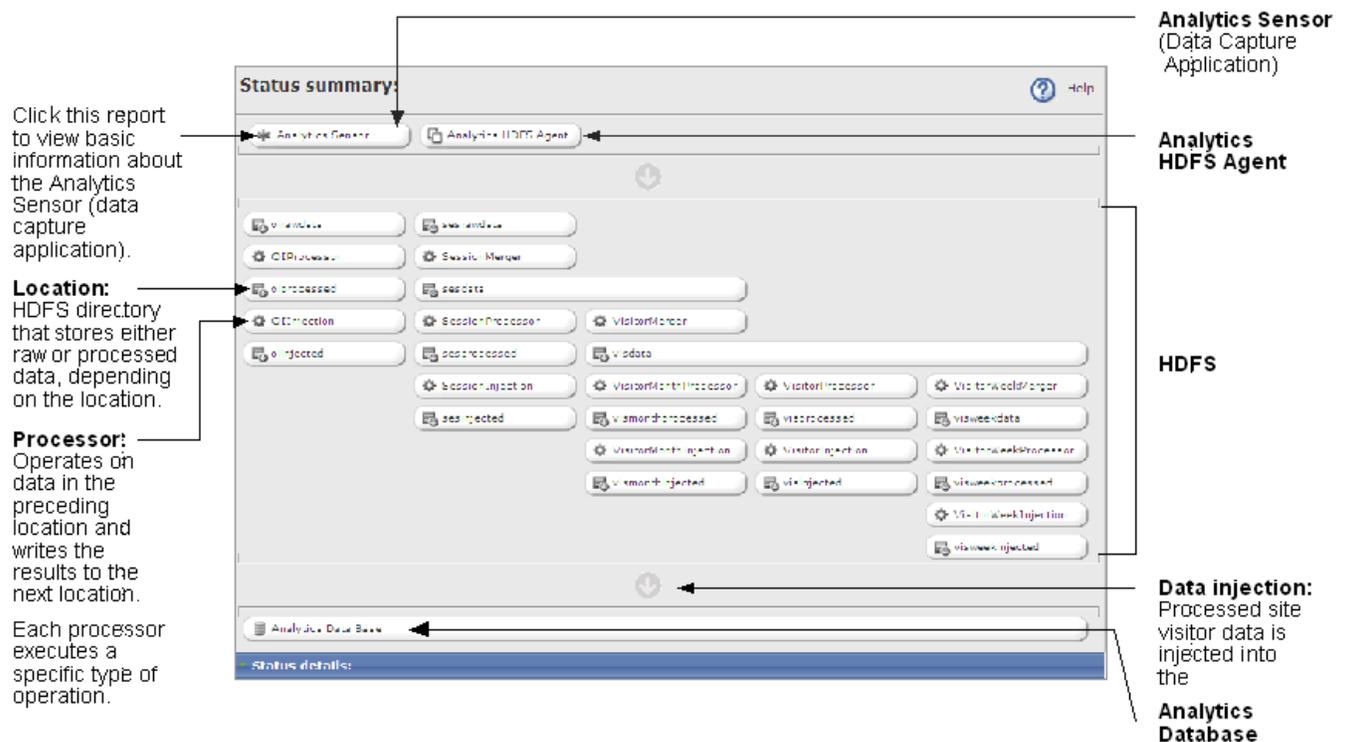
3.2 Monitoring Hadoop Jobs

Hadoop jobs can be monitored from the Status Summary panel of the Analytics Administration interface. The Status Summary panel renders an interactive flow chart which displays Hadoop job components—locations that store site visitor data captured by the Analytics Sensor (data capture application) and processors that calculate daily, weekly, and monthly sums for the stored site visitor data.

Accessing the Status Summary Panel

When working in the Analytics Administration interface, you can access the Status Summary panel by clicking the **Components** tab and selecting the **Overview** option (Figure 3–2).

Figure 3–2 Status Summary Panel (In the Components tab, the Overview option)



Each location stores different types of site visitor data. The type of site visitor data that is stored in a given location is determined by how that data is aggregated by the location's associated processor. For example, the `oiprocessed` location is associated with the `OIPProcessor` (it stores the results of the `OIPProcessor`'s computation) and therefore stores data such as the number of times specific assets have been rendered during a given time interval on a given date.

- Clicking a location enables you to view the status of the location and its data.
- Clicking a processor enables you to view the status of the data processing job.
- Clicking the **Analytics Sensor** and **HDFS Agent** buttons provides a status summary of those components. For more information about monitoring the Analytics Sensor, see [Section 2.1, "Sensor Overload Alerts."](#)

3.3 Processors and Locations

This section describes the different locations that are involved in storing site visitor data, and the processors that read the data from their locations, map/reduce the data, and write the results to another location.

- **Processors** – HDFS includes several processors developed by Oracle to process Analytics data. A processor consists of two parts: a mapper and a reducer. The mapper starts with a set of object impressions (collection of raw data) and creates intermediate data (n Java beans). The intermediate data is processed by the reducer in a way that aggregates the n Java beans into one Java bean containing x occurrences of a given data type, a second Java bean containing y occurrences of a different data type, and so on. As the reducer runs, it writes the aggregated data to the next location. The output of a processor is called a work package (for more information, see [Section 3.4.2, "Work Packages"](#)).

- [VisitorMerger](#)
- [VisitorMonthInjection](#)
- [VisitorMonthProcessor](#)
- [VisitorProcessor](#)
- [VisitorWeekInjection](#)
- [VisitorWeekMerger](#)
- [VisitorWeekProcessor](#)

The different locations that can be monitored from the Analytics Status Summary panel are listed alphabetically as follows:

- [oiinjected](#)
- [oiprocessed](#)
- [oirawdata](#)
- [sesdata](#)
- [sesinjected](#)
- [sesprocessed](#)
- [sesrawdata](#)
- [visdata](#)
- [visinjected](#)
- [vismonthinjected](#)
- [vismonthprocessed](#)
- [visprocessed](#)
- [visweekdata](#)
- [visweekinjected](#)
- [visweekprocessed](#)

3.4 Object Impressions and Work Packages

Object impressions and work packages are the main constructs of Hadoop Jobs. Object impressions are raw site visitor data that is captured as visitors browse and then processed by Hadoop Jobs in units called work packages. Results of the processing are stored to the Analytics database, where they are available on demand for the reports users generate.

This section contains the following topics:

- [Section 3.4.1, "Object Impressions"](#)
- [Section 3.4.2, "Work Packages"](#)

3.4.1 Object Impressions

An object impression is a single invocation of the `sensor` servlet. An object impression can also be thought of as a snapshot of raw site visitor data that is captured for analysis.

An object impression contains many types of raw data on the site visitor at the moment of capture. It contains session data and visitor data including object types, object IDs, sessions, session IDs, IP addresses, operating systems used, browsers used, referrers, and so on. If Engage is installed, the raw data also includes segments and recommendations.

When site visitors browse, object impressions are collected. They are collected during a 24-hour period as work packages in the oirawdata and sesrawdata locations.

3.4.2 Work Packages

A work package is a directory within a location, as shown in [Figure 3-4](#).

Figure 3-4 Work Package

<hadoop.hdfs.defaultfs>

```
[root@localhost bin]# tree -A /analytics/data/
/analytics/data/
├── oirawdata
│   ├── 2009
│   │   ├── 01
│   │   │   └── 14
│   │   │       └── Sensor-localhost.localdomain-10-1231924822209
│   │   │           └── data.txt
```

A work package stores:

- A data file containing either object impressions (raw data) or intermediate data (Java beans). The contents of the data file are statistically analyzed by a series of processors. When analysis is complete, the final processor injects results into the Analytics database for report generation.
- A metadata file. The metadata file reports the data processing status.

At least one work package exists in each location in the Hadoop Distributed File System (the number of work packages depends on the location. For an example, see [Section 3.4.2.1, "Data Collection"](#)). Each work package is positioned hierarchically in the location's directory structure, according to a calendar type of structure.

The data file in a given work package is input for the processor that is associated with the location containing the work package. When the processor completes its analysis of the data file, it writes the results, as a work package, to the next location for pickup by the next processor.

Note: The initial work package, containing newly captured object impressions, is created by the Analytics Sensor. All other work packages are created by the processors.

During data processing, neither work packages nor their contents are moved from one location to another. Instead, each work package's data file is read by the appropriate processor and analyzed by the processor. Results are written (by the processor) as a work package to the next location.

This section contains the following topics:

- [Section 3.4.2.1, "Data Collection"](#)
- [Section 3.4.2.2, "Processed Data"](#)
- [Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data"](#)
- [Section 3.4.2.4, "Directory Structure for Daily Work Packages"](#)
- [Section 3.4.2.5, "Directory Structure for Weekly Work Packages"](#)
- [Section 3.4.2.6, "Directory Structure for Monthly Work Packages"](#)

3.4.2.1 Data Collection

Object impressions are collected as work packages for a 24-hour period into two locations, simultaneously—`oirawdata` and `sesrawdata`. All work packages in the two locations contain a data file named `data.txt`. The locations (and their work packages) differ as follows:

- The `oirawdata` location collects objects impressions at fixed intervals during a 24-hour period; each interval has its own work package. The interval is specified by the `sensor.thresholdtime` property in the sensor's `global.xml` file. For example, if `sensor.thresholdtime` is set to 4 hours, then six work packages will have been collected in the `oirawdata` location at the end of 24 hours. All six packages are stamped with the creation time, and they all contain a `data.txt` file.
- The `sesrawdata` location collects object impressions continuously as a single work package during a 24-hour period. The work package is stamped with its creation time and contains a `data.txt` file.

Any one of the work packages in the `oirawdata` location contains only a portion of the day's raw data. A work package in the `sesrawdata` location contains the complete set of the day's raw data. In both locations, each work package is analyzed as soon as it is complete and computational resources are available.

3.4.2.2 Processed Data

All work packages are collected for a 24-hour period (see [Section 3.4.2.1, "Data Collection"](#)). The work packages are processed on a daily basis. For visitor data, additional work packages are created to represent weekly and monthly statistics. The work package directory structure for weekly and monthly processing differs from the directory structures for daily processing and data collection.

Note: Analytics administrators can obtain the directory structures of locations and paths to work packages from the HDFS file browser:

```
http://<hostname_MasterNode>:50070/
```

3.4.2.3 Directory Structure for Raw Data and oiprocessed Data

Work packages that contain raw data and `oiprocessed` data are stored in directories with a structure that identifies the day and time that the work packages were created. The following locations use a day-time directory structure: `oirawdata`, `sesrawdata`, and `oiprocessed` (an exception, as this location contains processed data).

The path to a raw data or `oiprocessed` work package is the following:

```
/<hadoop.hdfs.defaultfs>/<location>/<yyyy>/<mm>/<dd>/  
<workpackageDir>-<n>-<time>/data.txt <or part-<xxxxxx>
```

For example:

```
/analytics/data/oirawdata/2009/01/14/Sensor-localhost.localdomain-10-1231924822209
/data.txt
```

The variables are defined as follows:

- `<hadoop.hdfs.defaultfs>` is the location of the root directory under which raw data, output, and cache files are stored on the Hadoop file system.
- `<location>` is the name of the location that stores the raw data work package(s). Valid values for `<location>` are the following:
 - `oirawdata` collects data into multiple work packages during a 24-hour period
 - `sesrawdata` collects data into a single work packages during a 24-hour period
 - `oiprocessed` (although this location contains processed data)

(For a list of locations and descriptions, see [Section 3.3, "Processors and Locations."](#))

- `<yyyy>` is the year in which the work package was created.
- `<mm>` is the month in which the work package was created.
- `<dd>` is the day on which the work package was created. The day is determined from the site's time zone.
- `<workpackageDir>` is the sensor name (the IP address or host name of the data capture server)
- `<n>` is a system-generated number
- `<time>` is the creation time of the work package, computed in milliseconds elapsed since January 1, 1970.
- `data.txt` is the file that contains object impression. Raw data in the object impressions will be statistically analyzed by the processor that reads the file. `Data.txt` files are stored in the `oirawdata` location and in the `sesrawdata` location, as explained in [Section 3.4.2.1, "Data Collection."](#)

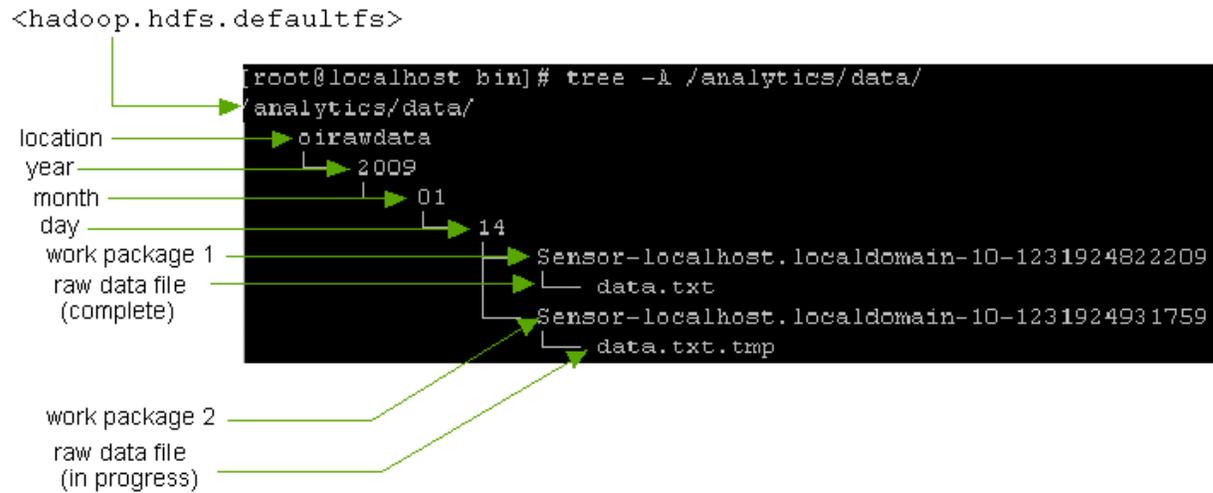
Note that all raw data files are named `data.txt`. The `<time>` stamp in the work package directory containing the data file uniquely identifies the data file.

- `<part-xxxxx>` is the name of the work package in the `oiprocessed` location.

[Figure 3-5](#) illustrates the directory structure of the `oirawdata` location. The first `data.txt` file that was created in the `oirawdata` location on January 14, 2009 is stored as shown in [Figure 3-5](#):

```
/analytics/data/oirawdata/2009/01/14/Sensor-localhost.localdomain-10-1231924822209
/data.txt
```

The second data file `data.txt.temp` is in progress. (At the end of the collection interval, the file will be complete and will take the name `data.txt`.) The file is stored as shown in [Figure 3-5](#).

Figure 3–5 oirawdata directory structure

3.4.2.4 Directory Structure for Daily Work Packages

When a raw data work package is complete, the associated processor statistically analyzes the work package's `data.txt` file and writes the results to the work package in the next location for pickup by the next processor.

Work packages that contain daily statistics are stored in directories with a structure that identifies the day on which the work package was created. The following locations use a day-based directory structure: `oiinjected`, `sesdata`, `sesprocessed`, `sesinjected`, `visdata`, `visprocessed`, and `visinjected`.

The path to a daily work package is the following:

```
<hadoop.hdfs.defaultfs>/<location>/<yyyy>/<mm>/<dd>/<workpackageID>/part-<xxxxxx>
```

For example:

```
/analytics/sesprocessed/2009/06/25/181bd6cd-c040-46a2-abb4/part-00000
```

The variables are defined as follows:

- `<location>` is the name of the location that stores the daily work package(s). Valid values for `<location>` are the following:
 - `oiinjected`, `sesdata`, `sesprocessed`, `sesinjected`, `visdata`, `visprocessed`, and `visinjected`
- `<workpackageID>` is a system-generated number used to identify the work package.
- The remaining variables are defined in [Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data."](#)

3.4.2.5 Directory Structure for Weekly Work Packages

Work packages that are processed for weekly statistics are stored in directories with a structure that identifies the ISO week in which the work package was stored. The following locations use a week-based directory structure: `visweekdata`, `visweekprocessed`, and `visweekinjected`.

The path to a weekly work package is the following:

- `visweekdata`

```
/<hadoop.hdfs.defaultfs>/visweekdata/<yyyy>/W<no.>/<yyyy>/<mm>/<dd>/<workpackag  
eID>/part-<xxxxx>
```

For example:

```
/analytics/visweekdata/2009/W26/2009/06/25/1db1039-0b10-417d-9895/part-00000
```

The variables are defined as follows:

- W<no.> represents the number of the week in the given year.
 - <workpackageID> is a system generated number.
 - The remaining variables are defined in [Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data."](#)
- visweekprocessed and visweekinjected

```
/<hadoop.hdfs.defaultfs>/<location>/<yyyy>/W<no.>/  
<workpackageID>/part-<xxxxx>
```

For example:

```
/analytics/visweekprocessed/2009/W26/9fe7607b-31b1-417d-9895/part-00000
```

The variables are defined as for visweekdata.

3.4.2.6 Directory Structure for Monthly Work Packages

Work packages that are processed for monthly statistics are stored in directories with a structure that identifies the month in which the work package was stored. The following locations use a month-based directory structure: vismonthprocessed and vismonthinjected.

The path to a monthly work package is the following:

```
/<hadoop.hdfs.defaultfs>/<location>/<yyyy>/<mm>/ <workpackageID>/part-<xxxxx>
```

For example:

```
/analytics/vismonthprocessed/2009/06/c3b9ex84-0417-4b6f-9e38/part-00000
```

The variables are defined as follows:

- <workpackageID> is a system-generated number used to identify the work package
- The remaining variables are defined in [Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data."](#)

3.5 Processor Descriptions

Analytics supports three types of processors. They analyze the same object impressions, collected within a 24-hour period, but they perform their computations differently.

This section contains the following topics:

- [Section 3.5.1, "Object Impression Processors"](#)
- [Section 3.5.2, "Session Data Processors"](#)
- [Section 3.5.3, "Visitor Data Processors"](#)

3.5.1 Object Impression Processors

Object Impression Processors analyze object impressions directly, by computing the frequency of occurrence of each type of data within the object impressions.

This section contains the following topics:

- [Section 3.5.1.1, "OIPProcessor"](#)
- [Section 3.5.1.2, "OIInjection"](#)

3.5.1.1 OIPProcessor

Output: Intermediate daily sums. This processor reads each work package that is created in the `oirawdata` location and computes an intermediate daily sum (that is, frequency of occurrence) for all types of data within the object impression.

Daily sums are called *intermediate* when they are computed for a work package containing less than 24 hours of data. Work packages are collected into the `oirawdata` location throughout the day, at the interval specified by the `sensor.thresholdtime` property in the sensor's `global.xml` file (for example, every 4 hours). Each work package then holds data that was collected for the specified interval—4 hours, in our example. At the end of 24 hours, six work packages will have been collected in the `oirawdata` location.

Details of the computation process are described below:

1. When a work package collected in the `oirawdata` location is complete, `OIPProcessor` reads the data file in the work package and counts (that is, sums, aggregates) the number of occurrences of each selected type of raw data in the work package. Thus, a work package has a set of intermediate daily sums, one for each selected type of raw data in the package. (If six work packages are collected and processed over a 24 hour period, then each work package has its own set of intermediate daily sums.)
2. `OIPProcessor` writes the intermediate daily sums for each work package to the `oiprocessed` location for pickup by the `OIInjection` processor.

Intermediate daily sums written to the `oiprocessed` location are counted (that is, summed, aggregated) by the `OIInjection` processor and injected into the Analytics database. The sum of intermediate daily sums for a given type of raw data is the grand total for the day for that type of raw data; it is called the complete daily sum, or aggregated daily sum.

Note: The WebCenter Sites database and Analytics database are not synchronized. Therefore, Analytics creates an `L2ObjectBean` object for each unique object impression. The `L2ObjectBean` saves the object impression's name (title) and object (asset) id in the `L2_Object` table of the Analytics database.

3.5.1.1.1 Input Location for OIPProcessor

oirawdata Stores the current day's `data.txt` file (and a metadata file). More information about `data.txt` can be found in [Section 3.4.2.1, "Data Collection."](#)

Table 3–1 Input Location for OIProcessor

oirawdata	Description
Directory structure	See Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data."
Work package	<workpackageDir>-<n>-<time>/data.txt
Work package data file	Contains beans of type <code>TransferObject</code> .
Data source	oirawdata folder on the local file system. The oirawdata location in the Hadoop File System is a duplicate of the oirawdata directory on the server where the Analytics Sensor is installed. Every 10 minutes (or the time interval that is explicitly set for the <code>sensor.thresholdtime</code> property), the Analytics Sensor creates a new work package (which is the input for the <code>OIProcessor</code>). Every object impression captured by the Analytics Sensor results in one line of data in the work package. The <code>/<dd></code> folder (see Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data.") represents the day of the object impressions (the day the work package was created).
Work package used by	This processor.

3.5.1.1.2 Output Location for OIProcessor

oiprocessed Stores the work packages of this processor. Each work package contains a data file with intermediate daily sums for the work package, that is, the frequency of occurrence of each type of data that was collected into the work package. (Each work package also contains a metadata file.)

Table 3–2 Output Location for OIProcessor

oiprocessed	Description
Directory structure	See Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data."
Work package	<workpackageDir>-<n>-<time>/part-<xxxxx>
Work package data file	Contains beans of type <code>L2ObjectBean</code> , <code>L3ObjecttypeBean</code> , <code>L3DownloadBean</code> , <code>L3InternalSearchBean</code> , and <code>L3ObjectBean</code> (and custom pre-aggregated object impression data).
Data source	This processor.
Work package used by	OIInjection processor.

3.5.1.2 OIInjection

Output: Complete daily sums for specific types of data (that is, the frequency of occurrence of each type of data that was collected during the last 24 hours). Injection status report.

1. This processor reads the intermediate daily sums in the data files of work packages in the `oiprocessed` location, and counts (that is, sums, aggregates) the intermediate daily sums. The result is a grand total — a complete daily sum for each type of data that was collected during the last 24-hour period.

2. This processor injects the complete daily sums into various tables in the Analytics database, and creates a status report in the `oiinjected` location. (For information about intermediate daily sums, see [OIProcessor](#).)

Data injected into the database is retrieved into the reports that Analytics users generate.

3.5.1.2.1 Input Location for OIInjection Processor. See [oiprocessed](#) ([Output Location for OIProcessor](#)).

3.5.1.2.2 Output Locations for OIInjection Processor

Analytics database Stores the output of this processor. The output is *complete* daily sums (that is, the frequency of occurrence of each type of data that was collected during the last 24- hours).

oiinjected Stores the work package created by this processor. Note that the work package does not contain a data file. It contains only the metadata file indicating the status of the injection process.

Table 3–3 Output Locations for OIInjection Processor

oiinjected	Description
Directory structure	See Section 3.4.2.4, "Directory Structure for Daily Work Packages."
Work package	<workpackageID>/part-<xxxxxx>
Work package data file	No data file is created in this work package. The metadata file (<code>.txt</code>) reports the status of the injection process.
Data source	This processor.
Work package used by	None of the processors. Administrators can open the HDFS file browser to view the metadata file (injection status report).

3.5.2 Session Data Processors

Session Data Processors analyze session objects derived from the object impressions.

This section contains the following topics:

- [Section 3.5.2.1, "SessionMerger"](#)
- [Section 3.5.2.2, "SessionProcessor"](#)
- [Section 3.5.2.3, "SessionInjection"](#)

3.5.2.1 SessionMerger

Output: Session objects for the last 24 hours of session data (that is, aggregated object impressions grouped by their respective sessions and stored in the `sesdata` location).

This processor reads the object impressions in the data file of the work package in the `sesrawdata` location. It takes session data from the object impressions and combines the data to create a session object for each entire session. The session object contains all the information that relates to the specific session. In this manner, `SessionMerger` aggregates all object impressions collected during a 24-hour period into their respective sessions. This processor writes the aggregated data (as a work package) to the `sesdata` location (for pickup by the [SessionProcessor](#)).

3.5.2.1.1 Input Location for SessionMerger Processor

sesrawdata Stores the current day's data.txt file (and a metadata file). More information about data.txt can be found in [Section 3.4.2.1, "Data Collection."](#)

Table 3–4 Input Location for SessionMerger Processor

sesrawdata	Description
Directory structure	See Section 3.4.2.3, "Directory Structure for Raw Data and oiprocessed Data." (Contains one work package for each calendar day.)
Work package	<workpackageDir>-<n>-<time>/data.txt
Work package data file	Contains beans of type RawSensorCallBean.
Data source	Analytics Sensor. The Analytics Sensor creates a new work package every 24 hours. (A work package contains the 24-hour interval of collected raw data, which is the input for the SessionMerger processor.) The work package folder in the directory structure represents the day of all the collected session information (the day is determined by the site's time zone).
Work package used by	SessionMerger processor.

3.5.2.1.2 Output Location for SessionMerger Processor

sesdata Stores the work package of the SessionMerger processor. The work package's data file contains session objects for the last 24 hours of session data (that is, aggregated object impressions grouped by their respective sessions). (The work package's metadata file contains the data processing status report.)

Table 3–5 Output Location for SessionMerger Processor

sesdata	Description
Directory structure	See Section 3.4.2.4, "Directory Structure for Daily Work Packages." (Contains one work package for each calendar day.)
Work package	<workpackageID>/part-<xxxxx>
Work package data file	Contains beans of type SessionBean.
Data source	This processor.
Work package used by	SessionProcessor .

3.5.2.2 SessionProcessor

Output: Complete daily sums of session data (that is, frequency of occurrence of each type of data across sessions that ran during the last 24 hours).

This processor reads the session objects in the data file of the work package in the sesdata location, computes complete daily sums, and writes the results (as a work package) to the sesprocessed location for pickup by the [SessionInjection](#) processor.

3.5.2.2.1 Input Location for SessionProcessor. See [sesdata \(Output Location for SessionMerger Processor\)](#).

3.5.2.2 Output Location for SessionProcessor

sesprocessed Stores the work package created by this processor. The work package's data file contains complete daily sums of session data (that is, the frequency of occurrence of each type of data across sessions that ran in the last 24-hours). (The work package's metadata file contains the data processing status report.)

Table 3–6 Output Location for SessionProcessor

sesprocessed	Description
Directory structure	See Section 3.4.2.4, "Directory Structure for Daily Work Packages." (Contains one work package for each calendar day.)
Work package	<workpackageID>/part-<xxxxx>
Work package data file	Contains beans of all L3* types, such as L3BrowserBean, L3ClickStreamBean, and L3SessionEntryBean (and custom pre-aggregated data).
Data source	This processor.
Work package used by	SessionInjection processor.

3.5.2.3 SessionInjection

Output: Injection status report.

This processor reads the complete daily sums in the data file of the work package in the `sesprocessed` location, injects the complete daily sums into various tables in the Analytics database, and creates a status report in the `sesinjected` location. Data injected into the database is retrieved into the reports that Analytics users generate.

3.5.2.3.1 Input Location for SessionInjection Processor. See [Section 3.5.2.2.2, "Output Location for SessionProcessor."](#)

3.5.2.3.2 Output Locations for SessionInjection Processor

Analytics database Stores complete daily sums of session data (that is, frequency of occurrence of each type of data across sessions that ran in the last 24 hours).

sesinjected Stores the work package created by this processor. Note that the work package does not contain a data file. It contains only the metadata file indicating the status of the injection process.

Table 3–7 Output Locations for SessionInjection Processor

sesinjected	Description
Directory structure	See Section 3.4.2.4, "Directory Structure for Daily Work Packages." (Contains one work package for each calendar day.)
Work package	<workpackageID>/part-<xxxxx>
Work package data file	No data file is created in this work package. The metadata file (<code>.txt</code>) reports the status of the injection process.
Data source	This processor.
Work package used by	None of the processors. Administrators can open the HDFS file browser to view the metadata file (injection status report).

3.5.3 Visitor Data Processors

Visitor data identifies site visitors by their IP addresses, for example. Visitor data includes the segments visitors belong to and the recommendations associated with those segments.

This section contains the following topics:

- [Section 3.5.3.1, "VisitorMerger"](#)
- [Section 3.5.3.2, "VisitorMonthProcessor"](#)
- [Section 3.5.3.3, "VisitorMonthInjection"](#)
- [Section 3.5.3.4, "VisitorProcessor"](#)
- [Section 3.5.3.5, "VisitorInjection"](#)
- [Section 3.5.3.6, "VisitorWeekMerger"](#)
- [Section 3.5.3.7, "VisitorWeekProcessor"](#)
- [Section 3.5.3.8, "VisitorWeekInjection"](#)

3.5.3.1 VisitorMerger

Output: Raw site visitor data.

This processor reads the visitor-specific data (such as segments and recommendations) from the data file in the work package of the `sesdata` location. It writes that visitor data (as a work package) to the `visdata` location, in raw format (not aggregated) in order to save all visitor IDs. The visitor data is not aggregated by this processor because it must be used in its raw form by other visitor data processors to compute daily, weekly, and monthly sums.

3.5.3.1.1 Input Location for VisitorMerger Processor. See `sesdata` ([Output Location for SessionMerger Processor](#)).

3.5.3.1.2 Output Location for VisitorMerger Processor

visdata Stores the work package created by this processor. The work package contains a data file with site visitor data in raw format. (The work package's metadata file contains the data processing status report.)

Table 3–8 Output Location for VisitorMerger Processor

visdata	Description
Directory structure	See 3.4.2.4, "Directory Structure for Daily Work Packages"
Work package	<workpackageID>/part-<xxxxx>
Work package data file	Contains beans of type <code>VisitorLocationBean</code> , <code>VisitorBean</code> (and custom visitor-related data).
Data source	This processor.
Work package used by	VisitorProcessor VisitorWeekMerger processor VisitorMonthProcessor

3.5.3.2 VisitorMonthProcessor

Output: Complete monthly sums for visitor data (that is, the frequency of occurrence of each type of visitor data that was collected during the last month).

This processor reads the raw visitor data in the data file of the work package in the `visdata` location and computes monthly sums. It writes the monthly sums (as a work package) to the `vismonthprocessed` location for pickup by the [VisitorMonthInjection](#) processor.

3.5.3.2.1 Input Location for VisitorMonthProcessor. See [visdata](#) ([Output Location for VisitorMerger Processor](#)).

3.5.3.2.2 Output Location for VisitorMonthProcessor

vismonthprocessed Stores the work package created by this processor. The work package's data file contains complete monthly sums for visitor data. (The work package's metadata file contains the data processing status report.)

Table 3–9 Output Location for VisitorMonthProcessor

vismonthprocessed	Description
Directory structure	See Section 3.4.2.6, "Directory Structure for Monthly Work Packages."
Work package	<workpackageID>/part-<xxxxx>
Work package data file	Contains beans of type <code>L3CityBean</code> , <code>L3CountryBean</code> , <code>L3RegionBean</code> , <code>L3VisitorBean</code> (and custom visitor-related data).
Data source	This processor.
Work package used by	VisitorMonthInjection processor.

3.5.3.3 VisitorMonthInjection

Output: Injection status report.

This processor reads the complete monthly sums in the data file of the work package in the `vismonthprocessed` location, injects the complete monthly sums into the Analytics database, and creates a status report in the `vismonthinjected` location. Data injected into the database is retrieved into the reports that Analytics users generate.

3.5.3.3.1 Input Location for VisitorMonthInjection Processor. See [vismonthprocessed](#) ([Output Location for VisitorMonthProcessor](#)).

3.5.3.3.2 Output Locations for VisitorMonthInjection Processor

Analytics database Stores the data from this processor's input location.

vismonthinjected Stores the work package created by this processor. Note that the work package does not contain a data file. It contains only the metadata file indicating the status of the injection process.

Table 3–10 Output Locations for VisitorMonthInjection Processor

vismonthinjected	Description
Directory structure	See Section 3.4.2.6, "Directory Structure for Monthly Work Packages."
Work package	<workpackageID>/part-<xxxxx>
Work package data file	No data file is created in this work package. The metadata file (.txt) reports the status of the injection process.
Data source	This processor.
Work package used by	None of the processors. Administrators can open the HDFS file browser to view the metadata file (injection status report).

3.5.3.4 VisitorProcessor

Output: Complete daily sums for visitor data (that is, the frequency of occurrence of each type of visitor data that was collected during the last 24 hours).

This processor reads the raw visitor data in the data file of the work package in the `visdata` location. It then computes complete daily sums and writes the complete daily sums (as a work package) to the `visprocessed` location for pickup by the [VisitorInjection](#) processor.

3.5.3.4.1 Input Location for VisitorProcessor. See [visdata](#) ([Output Location for VisitorMerger Processor](#)).

3.5.3.4.2 Output Location for VisitorProcessor

visprocessed Stores the work package created by this processor. The work package's data file contains complete daily sums for visitor data. (The work package's metadata file contains the data processing status report.)

Table 3–11 Output Location for VisitorProcessor

visprocessed	Description
Directory structure	See Section 3.4.2.4, "Directory Structure for Daily Work Packages."
Work package	<workpackageID>/part-<xxxxx>
Work package data file	Contains beans of type <code>L3CityBean</code> , <code>L3CountryBean</code> , <code>L3RegionBean</code> , <code>L3VisitorBean</code> (and custom visitor-related data).
Data source	This processor.
Work package used by	VisitorInjection processor.

3.5.3.5 VisitorInjection

Output: Injection status report.

This processor reads the complete daily sums in the data file of the work package in the `visprocessed` location, injects the complete daily sums into various tables in the Analytics database, and creates a status report in the `visinjected` location. Data injected into the database is retrieved into the reports that Analytics users generate.

3.5.3.5.1 Input Location for VisitorInjection Processor. See [visprocessed](#).

3.5.3.5.2 Output Locations for VisitorInjection Processor

Analytics database Stores the data from this processor's input location.

visinjected Stores the work package created by this processor. Note that the work package does not contain a data file. It contains only the metadata file indicating the status of the injection process.

Table 3–12 Output Locations for VisitorInjection Processor

visinjected	Description
Directory structure	See Section 3.4.2.4, "Directory Structure for Daily Work Packages."
Work package	<workpackageID>/part-<xxxxx>
Work package data file	No data file is created in this work package. The metadata file (.txt) reports the status of the injection process.
Data source	This processor.
Work package used by	None of the processors. Administrators can open the HDFS file browser to view the metadata file (injection status report).

3.5.3.6 VisitorWeekMerger

Output: Raw site visitor data from the visdata location merged into a weekly folder.

This processor reads the raw visitor data in the data file of the work package in the visdata location. It merges the raw site visitor data into the appropriate ISO-week directory (in the processor's work package). This processor does not modify the data. This processor then writes its work package to the visweekdata location.

3.5.3.6.1 Input Location for VisitorWeekMerger Processor. See [visdata \(Output Location for VisitorMerger Processor\)](#).

3.5.3.6.2 Output Location for VisitorWeekMerger Processor

visweekdata Stores the work package created by this processor. The work package's data file contains raw site visitor data (from the visdata location) merged into a weekly directory. (The work package's metadata file contains the data processing status report.)

Table 3–13 Output Location for VisitorWeekMerger Processor

visweekdata	Description
Directory structure	See Section 3.4.2.5, "Directory Structure for Weekly Work Packages."
Work package	<workpackageID>/part-<xxxxx>
Work package data file	Contains beans of type <code>VisitorLocationBean</code> and <code>VisitorBean</code> (and custom visitor-related data).
Data source	This processor.
Work package used by	VisitorWeekProcessor VisitorMonthProcessor

3.5.3.7 VisitorWeekProcessor

Output: Complete weekly sums for site visitor data (that is, the frequency of occurrence of each type of visitor data that was captured in the last week.)

This processor reads the weekly raw data in the data file of the work package in the `visweekdata` location. It computes weekly sums, and writes the weekly sums (as a work package) to the `visweekprocessed` location.

3.5.3.7.1 Input Location for VisitorWeekProcessor. See [visweekdata \(Output Location for VisitorWeekMerger Processor\)](#).

3.5.3.7.2 Output Location for VisitorWeekProcessor

visweekprocessed Stores the work package created by this processor. The work package's data file contains weekly sums for site visitor data (that is, the frequency of occurrence of each type of site visitor data that was captured during the last week.) (The work package's metadata file contains the data processing status report.)

Table 3–14 Output Location for VisitorWeekProcessor

visweekprocessed	Description
Directory structure	See Section 3.4.2.5, "Directory Structure for Weekly Work Packages."
Work package	<workpackageID>/part-<xxxxx> Note: If the work package contains visitor data for the last week of the year, it will also contain visitor data for the new year if the week runs over to the new year.
Work package data file	Contains beans of type <code>L3CityBean</code> , <code>L3CountryBean</code> , <code>L3RegionBean</code> , <code>L3VisitorBean</code> (and custom visitor-related data).
Data source	This processor.
Work package used by	VisitorWeekInjection processor.

3.5.3.8 VisitorWeekInjection

Output Injection status report.

This processor reads the weekly sums in the data file of the work package in the `visweekprocessed` location, injects the weekly sums into the Analytics database, and creates a status report in the `visweekinjected` location. Data injected into the database is retrieved into the reports that Analytics users generate.

3.5.3.8.1 Input Location for VisitorWeekInjection Processor. See [Section 3.5.3.7.2, "Output Location for VisitorWeekProcessor."](#)

3.5.3.8.2 Output Locations for VisitorWeekInjection Processor

Stores the data from this processor's input location.

Analytics database Stores the data from this processor's input location.

visweekinjected Stores the work package created by this processor. Note that the work package does not contain a data file. It contains only the metadata file indicating the status of the injection process.

Table 3–15 Output Locations for VisitorWeekInjection Processor

visweekinjected	Description
Directory structure	See Section 3.4.2.5, "Directory Structure for Weekly Work Packages."

Table 3–15 (Cont.) Output Locations for VisitorWeekInjection Processor

visweekinjected	Description
Work package	<workpackageID>/part-<xxxxxx>
Work package data file	No data file is created in this work package. The metadata file (.txt) reports the status of the injection process.
Data source	This processor.
Work package used by	None of the processors. Administrators can open the HDFS file browser to view the metadata file (injection status report).

Glossary

Analytics Data Capture Application

Also referred to as the "sensor."

Asset Registration

Enabling report generation for assets. Because WebCenter Sites assets are specific to a WebCenter Sites installation, you must register their asset types with Analytics by assigning them to reports through the Analytics Administration interface. This enables Analytics to:

- Recognize WebCenter Sites asset types
- Configure report menu options in the "General Information" and "Content Information" report groups
- Generate reports on assets of the registered asset types

Data Capture

The process of recording each visitor's clicks and the associated information—the date and time of each click, the assets that are clicked, the IP address from which the clicks are issued, the site being visited, and so on. The information is captured in real time by the sensor servlet and recorded in a `data.txt.tmp` file on the local file system (local to the Analytics data capture application). The `data.txt.tmp` file will be rotated by the sensor to `data.txt` when either the threshold interval is reached (see the `sensor.threshold` property on `sensor.thresholdtime`), or the application server is restarted.

Analytics can capture data on the usage of WebCenter Sites assets and on their visitors only if published pages are tagged for data capture. In the case of Engage assets, the assets themselves must be tagged for data capture.

Hadoop Jobs

Runs jobs in a parallel and distributed fashion in order to efficiently compute statistics on the raw data that is stored in the Hadoop Distributed File System.

Hadoop implements a computational paradigm named `Map/Reduce`, which divides a large computation into smaller fragments of work, each of which may be executed on any node in the cluster. `Map/Reduce` requires a combination of `jar` files and classes, all of which are collected into a single `jar` file that is usually referred to as a "job" file. To execute a job, you submit it to a `JobTracker`. Hadoop Jobs then responds with the following actions:

- Schedules and submits the jobs to `JobTracker`.

- Processes raw data captured by the data capture application into statistical data and injects the statistics into the Analytics database.

(Hadoop provides a web interface to browse HDFS and to determine the status of the jobs.)

Hadoop jobs pre-calculate commonly requested site usage statistics (such as average number of requests for a piece of content per unit time) in order to shorten report generation time. Statistical computation is typically resource-intensive and time-consuming. Therefore, it is performed not on-the-fly, each time a report is generated, but in advance so that it can be available by the time it is needed. Thus, precalculated statistics are immediately available for retrieval into reports. Statistics include, for example:

- Current information, such as today's total hits to each site, visiting countries, total number of visits from a given country, types of browsers, and average session duration.

- Historical results, such as:

Daily, weekly, and monthly statistics—for example, the total number of requests for a given asset on a given site during a certain month in the reporting period. Yearly statistics—a histogram in the performance indicator indicating the frequency with which certain assets were accessed during each week of the past year.

How long a Hadoop job runs depends on a number of factors, including site activity within the latest data capture time frame, the cumulative volume of captured data, and the configuration of the Analytics application. When data analysis is complete, the resulting statistics are available, at any time, for report generation.

Integration

Integrating Analytics with your WebCenter Sites system means enabling report generation for asset types and users on your online site. Integration involves registering content management sites, WebCenter Sites users, and asset types with Analytics, configuring the [Pageview Object](#) (through the "Page Views" report), and granting users permissions to access reports through membership in the appropriate user groups. The steps necessary to accomplish these tasks are described in [Chapter 1, "Integrating Oracle WebCenter Sites: Analytics with Oracle Web Center Sites."](#)

Internal Search

A search performed by a visitor using the site's built in search engine. This search returns results from within the site's contents.

Object

An Analytics construct. The subject of a report.

When storing and processing information, Analytics uses objects, whereas WebCenter Sites uses assets and asset types. To allow Analytics to recognize a WebCenter Sites asset type and track assets of that type, administrators define an Analytics object in terms of a WebCenter Sites asset type. They do so by configuring an Analytics report for the object and assigning the desired asset type to that object. The process of configuring a report defines the underlying asset.

Note: A special instance of an object is the [Pageview Object](#), which administrators must configure (by configuring the "Page Views" Report) in order for reports in the "General Information" group to work.

The "Page Views" report supports multiple asset types.

Object Impression

A single invocation of the sensor servlet. For more information, see [Section 3.4, "Object Impressions and Work Packages."](#)

Page View

An Analytics construct. A group of one or more assets, whose asset types are enabled for tracking by the Analytics data capture application.

Asset types are enabled for tracking when they are defined in the [Pageview Object](#) and when published pages displaying those asset types are tagged with `AddAnalyticsImgTag` (data capture tag). For more information about tracking, see [Data Capture](#).

Pageview Object

A default Analytics object that you configure through the "Page Views" report to specify the type (or types) of assets Analytics will track. Configuring the `Pageview` object enables default reports that are based on the `Pageview` object.

The `Pageview` object is the basis for the "Page Views," "Site Information," and "Clickstream" reports, and thus it should be assigned asset types whose assets make the most sense (from the marketing standpoint) to be included in these reports.

A `Pageview` object can be assigned multiple asset types. The "Page Views" report will contain statistics on the usage of those asset types.

"Page Views" Report

A report, based on the [Pageview Object](#). The "Page Views" report displays statistics on Page View activity on your site.

Processed Data

Visitor activity data that has been processed by Hadoop Jobs into statistical data. When processing is complete, the data is injected into the Analytics database, where it is immediately available for the reports that users request from the Analytics reporting interface.

Raw Data

Unprocessed data describing visitor activity on the site, recorded during the [Data Capture](#) process and stored in the local file system for future processing. This is the data on which statistics are calculated by the Hadoop Jobs for display in reports. (This data cannot be directly used for report generation.)

Sensor

Also referred to as the "Analytics data capture application."

Site Registration

Identifying a WebCenter Sites content management site to Analytics in order to enable Analytics to track visitor activity on assets published from that site.

Statistical Data

See [Processed Data](#).

Work Package

A collection of object impressions. For more information, see [Section 3.4, "Object Impressions and Work Packages."](#)