Fine Tuning Usage Data Processing

In order to ensure qualitative and accurate data, it is ideal to go through the fine tuning process on your system. This section provides detailed instructions on how to fine tune your system.

CardioLog collects usage data for various event types, including views, visits, duration, search, and most other ways a user can interact with your site. The Usage Data Processing service component matches events to the corresponding item in your SharePoint tree. Any event associated with a URL that does *not* match any item in your tree (eg., A URL address which includes a list of custom parameters, or a URL which is not part of your monitored environments) is known as a Lost Event. Report results do not display lost events by default. They appear only after fine tuning the data.

Fine tuning the data includes the following steps:

1. Determining which data should not be collected and defining data collection filters.
2. Identifying Lost Events and defining URL address modifications.

It is recommended to perform this process about a month after the initial installation in the production environment, 2-3 weeks after an upgrade, and 2-3 weeks after adding any new monitored environment to CardioLog. Fine tuning should be performed by a user with a local administrator account on the CardioLog server and with a CardioLog Administrator role. Contact us for further assistance.

Contents

Identifying Lost Events

1. Execute the following SQL script against the CardioLog database to get a list of lost event URLs. This can be done for a specific date range by editing the timestamp in the SQL query:

```
Use CardioLog
GO

select SearchURL, count(SearchURL)
from tab_event_log (nolock)
where timestamp >= '2010-09-01 00:00:00' /* Edit start date (date format: YYYY-MM-DD) */
and timestamp <= '2010-10-01 00:00:00' /* Edit end date (date format: YYYY-MM-DD) */
and entityid ='00000000-0000-0000-0000-000000000000'
and eventtype = 0
and SearchURL not like '%/_layouts/%' /* Ignore SharePoint central administration pages */
group by SearchURL
order by count(SearchURL) desc
```

Query results example:

```
/* Example #1: URL with custom parameters */
http://www.intlock.com/Pages/default.aspx?department=sales

/* Example #2: Access from the internal WFE*/
http://websrv01/Pages/default.aspx

/* Example #3: Access from an insecure channel (for SSL monitored environments such as https://www.intlock.com/) */
http://www.intlock.com/

/* Example #4: External events (non-monitored environments) */
http://www.amazon.com/
```

2. Check if the URL matches a SharePoint tree item. If it does not match, identify one of the reasons for the missing URL:

- URL with custom parameters
- URL with an internal server name, instead of FQDN
- URL with a non-secure channel, instead of a secure channel and vice-versa
- URL from a non-monitored environment

3. Create URL Mappings based on Regular Expressions. Examples:

- URL with custom parameters:

  pattern: ***"aspx\?department=.*"***
  action: ***"aspx"***

- URL with an internal server name, instead of FQDN:

  pattern: ***"http://websrv01/(.*)"***
  action: ***"http://www.intlock.com/${1}"***

- URL with a non-secure channel, instead of a secure channel or vice-versa:

  pattern: ***"http://www.intlock.com/(.*)"***
  action: ***"https://www.intlock.com/${1}"***

- Data for URLs from a non-monitored environment (external) or administration pages (under /_layouts) can be seen in the "Page Views By URL" and "Unique Users by URL" reports.

Fixing Lost Events

1. Make sure that there is a full backup of the CardioLog database before you continue to the next step.

2. Fix your history usage data according to your URL Mappings. The following example replaces the internal server name with the FQDN (Edit the timestamp in the SQL query to a relevant date range for you).
Create a script based on this example to fix history data according for the URL Mapping you have created. Then execute it against the CardioLog database.

```
/* Example: Replace the internal server name to the portal name - http://websrv01/ > http://www.intlock.com/ */
Use CardioLog
GO
declare @top int
declare @startTime datetime
declare @endTime datetime
set @top = 10000
set @startTime = '2010-09-01' /* Edit the start date (date format: YYYY-MM-DD) */
set @endTime = GETDATE() + 1
select top 1 '1' from tab_event_log
while @@rowCOunt > 0
begin
print cast(@top as varchar(50))
;with a as (select top (@top) url, searchUrl, QueryString
from tab_event_log LG
where
eventtype in (0,1)
and Timestamp >= @startTime
and Timestamp < @endTime
and entityid = '00000000-0000-0000-0000-000000000000'
and SearchURL like 'http://websrv01/%'
)

update a
set QueryString = Url,
Url = substring(replace(url, 'http://websrv01/','http://www.intlock.com/'), 0, 1000),
SearchUrl = substring(replace(SearchUrl, 'http://websrv01/','http://www.intlock.com/'), 0, 400)
end
GO
```

3. Execute the following SQL script against the CardioLog database to map the lost events to their corresponding SharePoint tree item. This can be done for a specific date range editing the timestamp in the SQL query:

```
USE [CardioLog]
```

```
GO

DECLARE @RC int
DECLARE @startTime datetime
DECLARE @endTime datetime

set @startTime = '2010-09-01' /* Edit start date (date format: YYYY-MM-DD) */
set @endTime = GETDATE() + 1

EXECUTE @RC = [dbo].[stp_eventlog_fix_lost_events]
@startTime
,@endTime
GO
```

Automatic Fine Tuning

The **SharePoint Tree Automatic Fine Tuning** service maps SharePoint URLs to their corresponding object in your SharePoint tree automatically. All items can then be accessed using the Object Explorer. If users access a SharePoint website through different zones (public URLs), or SharePoint pages with custom parameters, this service will map the different URLs into a single corresponding SharePoint object, bypassing the need to manually create an entry for multiple variations of the same item in the URL Mappings module.

Each SharePoint object in the SharePoint tree structure has a unique SharePoint object ID, known as an SPID. The Portal Tree Updates service component retrieves all lost events and sends them to the SPIDFinder web service, located on your SharePoint WFEs, and then maps them by their SPID within the tree structure.

The SharePoint Tree Automatic Fine Tuning web service is installed with the CardioLog Analytics SharePoint feature by default.

To enable/disable the SharePoint Tree Automatic Fine Tuning web service, edit the [CardioLog Installation Folder]/CardioLogScheduleServices/CardioLog.Services.exe.config (located on the CardioLog application server):

```
<add key="RunSpidFinder" value="true"/>
Enable/disable the SharePoint Tree Auto Fine Tuning web service. Values - true/false

<add key="GetLostUrlsRegExExclude" value="_layouts|_vti_bin|editform\.aspx|newform\.aspx|aspx&amp;"/>
A regular expression for excluding URLs to be fine-tuned.
```

You can test the Web Service by browsing /_layouts/CardioLogAgent/SpidFinder.asmx and invoking the GetSPIDbyURL method. Submit the URL for your SharePoint website homepage, and verify that the guide value is not "00000000-0000-0000-0000-000000000000"

Refreshing Report Data

1. Clear the report data cache:
   From the navigation pane in CardioLog, go to Administration > System Configuration > Reporting Data and click on **Clear Cache**
2. Delete existing versions of the report by using the navigation pane in CardioLog, and clicking on Report Center. Click the relevant report and click on **Delete Historical Data**.
3. Regenerate reports. CardioLog reports are generated automatically by the Report Scheduling service. To regenerate a report on your own, click the report and select **Open**, then click **Generate Report** from the actions menu.