Oracle® Communications Offline Mediation Controller

Duplicate Check Cartridge User's Guide

Release 6.0

E64448-01

June 2015

ORACLE®

Oracle Communications Offline Mediation Controller Duplicate Check Cartridge User's Guide, Release 6.0

E64448-01

Contents

Preface

This document describes how to use the Oracle Communications Offline Mediation Controller duplicate check enhancement processor (EP) cartridge to handle duplicate call detail records (CDRs).

Audience

This document is intended for solution designers who configure Offline Mediation Controller.

Downloading Oracle Communications Documentation

Product documentation is located on Oracle Help Center:

http://docs.oracle.com

Additional Oracle Communications documentation is available from the Oracle software delivery Web site:

https://edelivery.oracle.com

Related Documents

For more information, see the following documents:

- *Offline Mediation Controller Cartridge Development Kit Developer's Guide*: For information about how to develop a cartridge.
- *Offline Mediation Controller Cartridge Development Kit NPL Reference Guide*: For information about how to use the Node Programming Language for developing or extending a cartridge.

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 that have purchased support 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.

1

Duplicate Check Cartridge Overview

This chapter provides an overview of the Oracle Communications Offline Meditation Controller duplicate check enhancement processor (EP) cartridge, which identifies duplicate records while processing call detail records (CDRs).

Before reading this chapter, you should be familiar with:

- Offline Mediation Controller cartridge concepts. For more information, see *Offline Mediation Controller Cartridge Development Kit Developer's Guide*.

About Duplicate Check

Offline Mediation Controller duplicate check is the process of identifying duplicate records while processing CDRs. The Offline Mediation Controller duplicate check EP cartridge creates partitions (in memory data structures) and duplicate check files to store duplicate check keys. The duplicate check keys are used to identify any duplicate records in the incoming CDRs.

Before you can use the duplicate check EP, do the following:

- Create and configure the Offline Mediation Controller duplicate check EP cartridge node. For more information, see "Creating and Configuring the Duplicate Check Cartridge Node".

About Partitions

When processing CDRs, the Offline Mediation Controller duplicate check EP node performs duplicate checks by storing the duplicate check keys in partitions and in duplicate check files. A duplicate check key is a set of fields in the CDR that uniquely identify the record and the record timestamp. The key is stored in the partition and in the duplicate check file for a predefined time period.

The duplicate check EP node creates the partitions as configured based on the CDR timestamp. The partitions are created for the configured time period. When a CDR is processed, the duplicate check node creates the partition in memory and adds the duplicate check keys to the partition and to the duplicate check file.

When subsequent CDRs are processed, the duplicate check node compares the keys in the CDR to the keys in the partition. If the key already exists in the partition, the record is rejected and is written to a file in the duplicate record storage directory, and the information related to the duplicate records is written to a log file. If the key does not exist, it is added to the partition and the duplicate check file.

Workflow for CDRs

When a duplicate check EP node is configured in the mediation node chain, depending on the configuration in the NPL rule file, the records are processed as follows:

1. The collection cartridge (CC) node processes the CDR input file.
2. If the input file is successfully processed by the CC node, the duplicate check EP node processes the CDR data as follows:
   - Memory partitions and duplicate check files are created based on the configured partition size and the CDR timestamp.
   - When a CDR arrives, the appropriate partition is loaded into memory.
   - The duplicate check keys are stored in the appropriate partition and the duplicate check file based on the CDR timestamp.
   - The key from the incoming CDRs belonging to the same time interval is compared with the keys stored in the partition.
   - If a duplicate key is found in the partition, the record is rejected and the duplicate record is written to a file in the duplicate record storage directory.
   - If a duplicate key is not found, the key is added to the list of keys in the partition and the duplicate check file, and the CDR is distributed to the next node in the mediation node chain.
   - If the maximum number of partitions to be stored in memory is reached, the oldest partition is deleted from memory.
   - When the partition reaches its configured retention time, the partition and the duplicate check file are deleted.

   For example:

   - If a duplicate check EP node is configured with the partition size set to **Hourly** and the retention time set to 24 hours:

     When a CDR arrives with the following data:

     ```
     20140723104450,9945168238,VOICE,101
     20140723105450,9945168239,VOICE,102
     20140723104050,9945168240,VOICE,103
     20140723103050,9945168241,VOICE,104
     ```

     a new partition is created for the hour with the timestamp **2014-07-23T10-00-00.000_2014-07-23T11-00-00.000**, the keys are stored in the partition and the duplicate check file, and the duplicate flag in the record is set to **0**, indicating it as a non-duplicate record.

     When another CDR arrives with the following data:

     ```
     20140723104450,9945168238,VOICE,101
     20140723105450,9945168239,VOICE,102
     ```

     the CDR data keys are compared with the keys stored in the partition, and the duplicate flag in the record is set to **1,** indicating it as a duplicate record.

     When another CDR arrives with the following data:

     ```
     20140722084450,9945168238,VOICE,101
     20140722085450,9945168239,VOICE,102
     ```

the CDR is identified as an old record, and the duplicate flag in the record is set to **-1**.

After 24 hours, the partition and the duplicate check file that have reached the retention time are deleted.

- If a duplicate check EP node is configured with the partition size set to **Daily** and the retention time set to 2 days:

  When a CDR arrives with the following data:

  ```
  20140723104450,9945168238,VOICE,101
  20140723105450,9945168239,VOICE,102
  20140723104050,9945168240,VOICE,103
  20140723103050,9945168241,VOICE,104
  ```

  a new partition is created for the day with the timestamp **2014-07-23T00-00-00.000_2014-07-24T00-00-00.000**, the keys are stored in the partition and the duplicate check file, and the duplicate flag in the record is set to **0**, indicating it as a non-duplicate record.

  When another CDR arrives with the following data:

  ```
  20140723104450,9945168238,VOICE,101
  20140723105450,9945168239,VOICE,102
  ```

  the CDR data keys are compared with the keys stored in the partition, and the duplicate flag in the record is set to **1**, indicating it as a duplicate record.

  When another CDR arrives with the following data:

  ```
  20140720104450,9945168238,VOICE,101
  20140720105450,9945168239,VOICE,102
  ```

  the CDR is identified as an old record, and the duplicate flag in the record is set to **-1**.

  After 2 days, the partition and the duplicate check file that have reached the retention time are deleted.

3. The next node in the mediation node chain processes the CDR, which is distributed to the target system.

2

Creating and Configuring the Duplicate Check Cartridge Node

This chapter describes how to create and configure the Oracle Communications Offline Mediation Controller duplicate check enhancement processor (EP) cartridge node to check for duplicate records while processing call detail records (CDRs).

Before reading this chapter, you should be familiar with:

- Offline Mediation Controller cartridge concepts and Node Programming Language (NPL).

About the Duplicate Check EP Cartridge Node

The duplicate check EP node creates the partitions as configured based on the CDR timestamp. The partitions are created for the configured time period. When a CDR is processed, the duplicate check node creates the partition in the memory and adds the duplicate check keys in the partition and in the duplicate check file.

When new CDRs are processed, the duplicate check node compares the keys in the CDR with the keys in the partition. If the key already exists, the record is rejected and written to a file in the duplicate record storage directory, and the information related to the duplicate records are written to a log file. If the key does not exist, it is added to the partition and the duplicate check file.

Creating a Duplicate Check EP Node

To create a duplicate check EP node:

1. Log on to Offline Mediation Controller Administration Client.

   The Node Hosts & Nodes (logical view) screen appears.
2. In the **Mediation Hosts** table, select a host.
3. In the **Nodes on Mediation Host** section, click **New**.

   The Create a Node dialog box appears.
4. Select **Wireless** and click **Next**.
5. Select **Enhancement Processor (EP)** and click **Next**.
6. Select **Duplicate Check EP** and click **Finish**.

   The New Node dialog box appears.
7. In the **Name** field, enter a name for the node.

8. From the **Rule File** list, select the rule file that matches the type of input file processed by the EP node.

   To edit the rule file, see "Configuring the NPL Rule File for Duplicate Check EP".

9. Click the **General** tab and do the following:

   a. From the **Debug** list, select one of the following:

      To log short debug messages in the node log file, select **OFF**.

      To log detailed debug messages in the node log file, select **ON**.

   b. In the **Max Log File Size** field, enter the maximum size in bytes for the log file. When the log file reaches its limit, the node closes the file and opens a new file. The minimum value is **50000** and the maximum value is **2000000000**.

   c. Select the **Enable Statistics** check box, which enables node statistics.

   d. Select the **Enable bulk read/write** check box, which enables the node to read or write records in bulk.

   e. In the **Read Timer** field, enter the number of seconds Offline Mediation Controller waits before checking for incoming records. The minimum value is **1** and the maximum value is **3600**.

   f. In the **NARs Per File** field, enter the maximum number of NARs allowed in an output file. The minimum value is **1** and the maximum value is **10000**.

   g. In the **Idle Write Time** field, enter the number of seconds Offline Mediation Controller waits before transferring the NAR output file to the output directory of the processing node, whether or not it has reached its maximum size. The minimum value is **1** and the maximum value is **3600**.

   h. Select the **Backup NAR Files** check box, which enables the node to back up each processed NAR file.

   i. If you selected **Backup NAR Files**, enter the number of days to retain the backup NAR files in the **NAR File Retention Period** field.

   j. Select the **Input Stream Monitoring** check box, which enables the node to monitor the input stream of records and trigger an alarm if no input records arrive for the set interval.

   k. If you selected **Input Stream Monitoring**, enter the duration of time that the node waits for the input stream of records before triggering an alarm if there is no input in the **Interval** field; select the time period: **Day**, **Hour**, or **Minute**.

10. Click the **Threading Options** tab and do the following:

    - In the **Number of Processor Threads** list, select the number of processing threads. The maximum value is **20**.

11. Click the **De-Duplication** tab and do the following:

    a. Do one of the following:

       To use a scratch directory to store the duplicate check files containing the keys, select the **Use Scratch Directory** check box.

       To use a different directory from the scratch directory to store the duplicate check files containing the keys, enter the path of the directory in the **Storage Directory** field.

    b. From the **Partition Size** list, select the type of partition: **Daily** or **Hourly**.

c. In the **Retention Limit** field, enter how long to retain the partition based on the type of partition. For example, if the type of partition is **Daily**, the value is in days.

d. In the **Number of Partitions in Memory** field, enter how many partitions to retain in memory.

e. In the **Duplicate Records Storage Directory** field, enter the path of the directory where all the files containing duplicate records are stored.

f. From the **Duplicate Records Push Time Unit** list, select the time unit for the time that the node waits before the file containing the duplicate records is moved from the scratch directory to the storage directory.

g. In the **Duplicate Records Push Time Period** field, enter the duration of time that the node waits before the file containing the duplicate records is moved from the scratch directory to the storage directory.

h. In the **Number of Duplicate Records per file** field, enter the maximum number of duplicate records to store in the file. The minimum value is **1** and the maximum value is **10000**. The default is **2000**.

12. Click the **Destination** tab and do the following:

a. Select the **Enable** check box, which enables the connection between the EP node and any destination cartridge node.

b. From the **Routing** list, select one of the following:

If the **Enable** check box is not selected, **Routing** is set to **None**.

To enable multicast routing between the EP node and the destination cartridge node, select **Multicast**.

To enable round-robin routing between the EP node and the destination cartridge node, select **Round Robin**.

13. Click **Save**.

Configuring the NPL Rule File for Duplicate Check EP

To configure the NPL rule file for duplicate check EP:

1. Log on to Offline Mediation Controller Administration Client.

The Node Hosts & Nodes (logical view) screen appears.

2. In the **Mediation Hosts** table, select the mediation host that contains the duplicate check EP node.

3. In the **Nodes on Mediation Host** section, select the duplicate check EP node that you want to configure, and click **Edit**.

The Node dialog box appears.

4. From the **Rule File** list, select one of the sample NPL rule files.

5. Click **Edit** for the selected rule file.

The NPL Editor dialog box appears.

6. In the configuration block, do the following:

- Add the following entry, which defines the CDR fields to be used as the duplicate check keys:

```
DUP_CHECK_KEYS "FieldName1 FieldName2 ...";
```

where *FieldNameX* is the name of the field in the CDR.

For example:

```
DUP_CHECK_KEYS "calling_number session_id seq_no";
```

- Add the following entry, which defines the event timestamp based on which the partition is created:

  DUP_CHECK_EVENT_TIME_FIELD "*TimestampField***";**

  where *TimestampField* is the name of the timestamp field in the CDR.

  For example:

  ```
  DUP_CHECK_EVENT_TIME_FIELD "start_time";
  ```

7. Compile and save the file.
8. Close the NPL Editor dialog box.
9. Click **Save**.

   The configuration is saved.