



Adding personal data identifiers using Data Fusion

Cloud Manager

Tom Onacki
December 14, 2020

Table of Contents

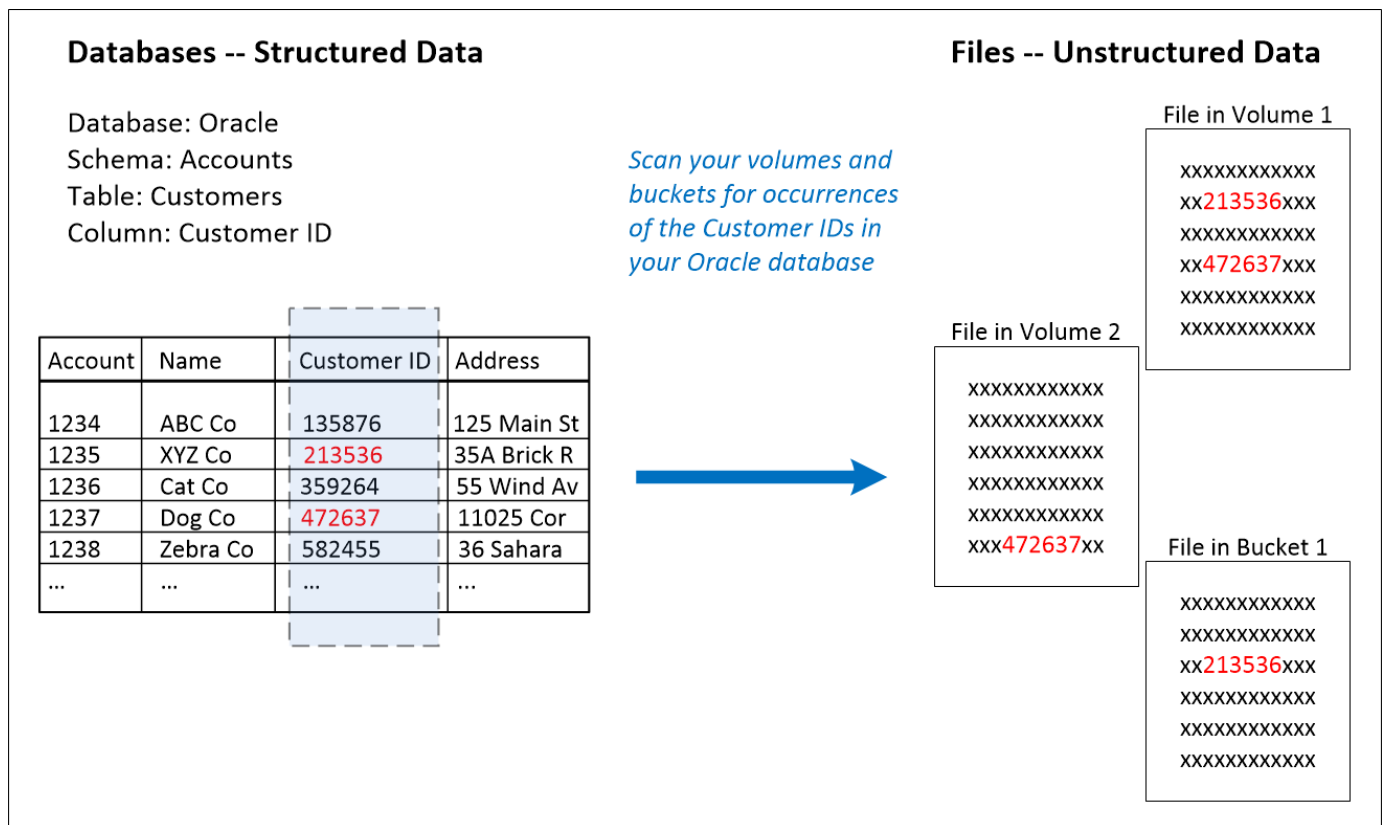
- Adding personal data identifiers using Data Fusion..... 1
 - Creating custom personal data identifiers from your databases 1
 - Deleting a Data Fusion source..... 2

Adding personal data identifiers using Data Fusion

A feature we call *Data Fusion* allows you to scan your organizations' data to identify whether unique identifiers from your databases are found in files or other databases - basically making your own list of "personal data" that is identified in Cloud Compliance scans. This gives you the full picture about where potentially sensitive data resides in *all* your files.

Creating custom personal data identifiers from your databases

You can choose the additional identifiers that Cloud Compliance will look for in its' scans by selecting a specific column, or columns, in a database table. For example, the diagram below shows how data fusion is used to scan your volumes, buckets, and databases for occurrences of all your Customer IDs from your Oracle database.

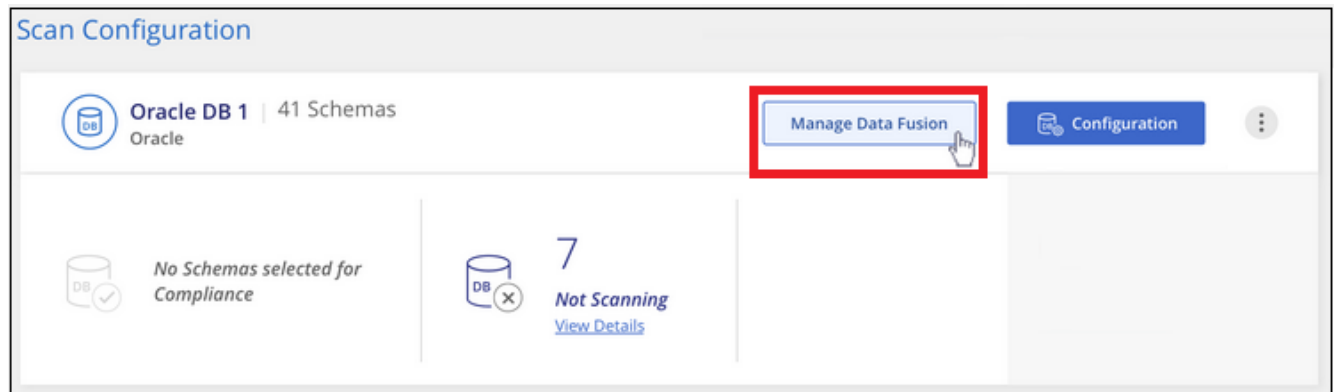


As you can see, two unique Customer IDs have been found in two volumes and in one S3 bucket. Any matches in database tables will also be identified.

Steps

You must have [added at least one database server](#) to Cloud Compliance before you can add data fusion sources.

1. In the Scan Configuration page, click **Manage Data Fusion** in the database where the source data resides.



2. Click **Add Data Fusion source** on the next page.
3. In the *Add Data Fusion Source* page:
 - a. Select the Database Schema from the drop-down menu.
 - b. Enter the Table name in that schema.
 - c. Enter the Column, or Columns, that contain the unique identifiers you want to use.

When adding multiple columns, enter each column name, or table view name, on a separate line.

4. Click **Add Data Fusion Source**.

The Data Fusion inventory page displays the database source columns that you have configured for Cloud Compliance to scan.

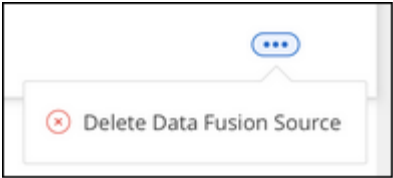
'DB Name 1' Data Fusion			+ Add Data Fusion source
With Data Fusion, Cloud Compliance can identify occurrences of your organization's unique identifiers found in your unstructured data stores, using structured data indexes containing those unique identifiers as a source reference. Learn More			
Database Schema	Table	Data Fusion Source Columns	
SchemaName1	Table 1	Column 12, Column 14, Column 18	...
SchemaName2	Table 2	Column 12, Column 14, Column 18	...

Results

After the next scan, the results will include this new information in the Dashboard under the "Personal" results section, and in the Investigation page in the "Personal Data" filter. Each source column you added appears in the filter list as "Table.Column", for example `Customers.Customer ID`.

Deleting a Data Fusion source

If at some point you decide not to scan your files using a certain Data Fusion source, you can select the source row from the Data Fusion inventory page and click **Delete Data Fusion Source**.



Copyright Information

Copyright © 2021 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system-without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.