Bulk insert data conversion error for varbinary data type

Last updated by | Vitor Tomaz | Feb 18, 2021 at 3:14 AM PST

Contents

Issue

The customer is trying to bulk insert from a txt file to a table and gets data conversion error (type mismatch or invalid character for the specified codepage) for a varbinary column.

Value sample: (0x08D290BB66ADF27D36E336443711A10785846F18).

Workaround:

When removing the hex prefix "0x" it works fine, and value inserted as it should be.

Internal Reference

Master ICM 218062042 ↗

If we have identified similar issue, please tag the support case after providing the work around. If Customer is looking for another work around, create a new ICM and tag support case and Link the Master ICM.

RCA:

This is by design behavior. Varbinary must be serialized without 0x part in csv file.

There is some work happening for Azure DW to enable reading binary data and 0x. Here is a piece of info regarding this from Miro Flasza: "I introduced a Polybase compatibility mode override in Openrowset and will be allowing reading binary with the 0x prefix and bit type as true/false when the override is enabled (working on those parsing changes this week). We saw this with a couple of other customers already. We can later decide, if we want to extend this functionality to BULK INSERT going forward."

So, BULK INSERT behavior is by design for now and later we will decide if Polybase compatibility mode should be added for BULK INSERT. Good thing is that once it is implemented and enabled for OPENROWSET, it should be small addition to enabled it for BULK INSERT

Root Cause Classification

TBD

How good have you found this content?