

[SAP ODP] SAP full parameter support

Last updated by | Zhangyi Yu | Feb 26, 2023 at 6:21 PM PST



Contents

- [Issue](#)
- [Root Cause](#)
- [Resolution](#)
 - [Case-1:](#)
 - [Case-2:](#)
- [Reference:](#)
- [Additional Information:](#)

Issue

Based on customers' feedback, ADF has supported full parameter on every SAP property/option exposed on UX. Customer can use parameter to set dynamic value to server extracting full/initial/delta data from different SAP tables.

Today ADF support two type parameter and customer may wrongly use them in dataflow SAP.

- Pipeline Expression Builder (<https://learn.microsoft.com/en-us/azure/data-factory/control-flow-expression-language-functions> )
 - JSON syntax/language
 - Support parameters used in ADF linked Service/dataset/activity/pipeline/IR/...
 - Related parameter value resolve is happened on ADF micro-services, when ADF compute (Azure IR, SHIR, SSIS-IR...) gets activity payload, the value of parameters are already plaintext.
 - This type parameter is most popular among customers when trigger/debug pipeline/activity execution.
- Dataflow Expression Builder (<https://learn.microsoft.com/en-us/azure/data-factory/concepts-data-flow-expression-builder> )
 - DSL syntax/language
 - Support parameters used in dataflow script.

```

"scriptlines": [
  "parameters(",
  "  dfContext as string,",
  "  dfObject as string,",
  "  dfKeyColumns as string[],",
  "  dfReadMode as string,",
  "  dfPartitions as string",
  ")",
  "source(allowSchemaDrift: true,",
  "  validateSchema: false,",
  "  store: 'SapOdp',",
  "  format: 'sapObject',",
  "  staged: true,",
  "  context: ($dfContext),",
  "  objectName: ($dfObject),",
  "  readMode: ($dfReadMode),",
  "  enableCdc: true,",
  "  keys: ($dfKeyColumns)) ~> source1",
]

```

- Related parameter value resolve is happened in dataflow runtime inside spark cluster
- This type parameter is mainly used to server customer have dataflow test-connection/import-projection/preview-data/... to verify parameter value(s) at authoring before having pipeline/activity execution.

Root Cause

Resolution

For support SAP as source with parameters in dataflow.

Name	Type	Default value
df_context	abc string	Enter expression...
df_objectName	abc string	Enter expression...
df_keyColumns	abc string[]	Enter expression...
df_readMode	abc string	Enter expression...
df_partitions	abc string	Enter expression...

Parameter	Value	Refresh
ODP context *	\$df_context	Refresh
ODP name *	\$df_objectName	Refresh
Run mode *	\$df_readMode	Refresh
Key columns *	\$df_keyColumns	Refresh






Note:

- The parameter type of sapContext/sapObjectName/sapRunMode is string.
- The parameter type of sapKeyColumns is array of string while its syntax is like ['x','y'], not (['1', '2']) which is custom expression. **This parameter is required and valid only when customer use run mode "Full on the first run, then incremental" or "Incremental changes only" to extract SAP data.**
- The parameter type of sapPartitions is string. **This parameter is required/valid only when customer use run mode "Full on every run" or "Full on the first run, then incremental" to extract SAP data.**

Case-1:

If customer wants to use parameter in dataflow SAP, we strongly suggest them to use existed UX template, then have few adjustments on sink store (sink is ADLS Gen2 + delta format by default, customer can feel free to change it to be snowflake/SQL DW/...).

The pattern of this template is "lookup + foreach + dataflow" (Multiple SAP tables' CDC can be easily handled in single pipeline), customer just need to provide one file/table/... to list each SAP table's settings.

UX Template	Lookup Expected Output
<p>Template gallery</p> <p> Azure Data Factory has a comprehensive knowledge c Learn more about our SAP Knowledge Center </p> <p> multiple tables from SAP OI  Categories : All</p> <div>  <p>Replicate multiple tables from SAP ODP to Azure Data Lake... by Microsoft</p> <p>Use this template to replicate multiple tables from SAP to Azure Data Lake Gen2 in Delta format with key partition.</p> </div>	<pre> { "count": 2, "value": [{ "checkPointKey": "cba2acf0-d5e2-4d84-a552-e0a059b6d320", "sapContext": "ABAP_CDS", "sapObjectName": "ZPERFCDPOSSF", "sapRunMode": "fullAndIncrementalLoad", "sapKeyColumns": ["TABKEY"], "sapPartitions": [[{ "fieldName": "TEXTCASE", "sign": "1", "option": "EQ", "low": "1" }, { "fieldName": "TEXTCASE", "sign": "1", "option": "EQ", "low": "X" }]], "deltaContainer": "delta", "deltaFolder": "ZPERFCDPOS", "deltaKeyColumns": ["TABKEY"], "deltaPartition": "TEXTCASE", "stagingStorageFolder": "stagingcontainer/stagingfolder" }] } </pre> <p>The source (SAP table-1) and sink settings</p> <p>Here only one partition is specified with multiple conditions. Customer can feel free to specify more partitions with same format</p> <pre> { "checkPointKey": "fgaeca7f-d3d4-406f-bb48-a17faa83f76c", "sapContext": "SAPI", "sapObjectName": "Z0131", "sapRunMode": "incrementalLoad", "sapKeyColumns": ["ID"], "sapPartitions": [], "deltaContainer": "delta", "deltaFolder": "Z0131", "deltaKeyColumns": ["ID"], "deltaPartition": "COMPANY", "stagingStorageFolder": "stagingcontainer/stagingfolder" } </pre> <p>The source (SAP table-2) and sink settings</p>

Case-2:

If customer wants to set parameter in dataflow reference, he must follow DSL syntax in dataflow expression builder. **If you set value in dataflow expression builder, mostly it is used to server dataflow debug (test connection, preview data...), or used as default value if there is no new value passed from pipeline level.**

- The parameter value of sap key columns is like:

Dataflow expression builder

Default value

```
['col-1', 'col-2']
```

- The parameter value of sap partitions:
 - Step-1: Use online tool <https://jstontstring.com/> to covert JSON (array of array of object) into string.
 - Step-2: Update start/end character of the string to be single quote, for example:

```
'[[{"fieldName":"LANGU","sign":"I","option":"EQ","low":"1"},{"fieldName":"DATATYPE","sign":"I","option":"EQ","low":"ACCP"}],[{"fieldName":"LANGU","sign":"I","option":"EQ","low":"2"},{"fieldName":"DATATYPE","sign":"I","option":"EQ","low":"CHAR"}]]'
```

- Step-3: Paste up content into dataflow expression builder and save it.

Dataflow expression builder

Default value

```
'[[{"fieldName":"LANGU","sign":"I","option":"EQ","low":"1"},{"fieldName":"DATATYPE"
```

Note: For same parameter used in dataflow reference, if customer has define its value at pipeline level + default value at dataflow, ADF will honor pipeline level value first.

Reference:

1. <https://learn.microsoft.com/en-us/azure/data-factory/solution-template-replicate-multiple-objects-sap-cdc>

Additional Information:

- Icm Reference: N/A
- Author: Zhangyi Yu
- Reviewer: Zhuoyang Zhang