SAP connection tools for process automation: Microsoft, Pentaho, Talend









CONTENT

1.	INTRODUCTION	
2.	PENTAHO	
3.	TALEND	
4.	POWER AUTOMATE	
5.	AZURE DATA FACTORY / SYNAPSE ANALYTICS	
6.	CONCLUSIONS	64



1. **INTRODUCTION**

The document gathers a series of tools used for process automation. A brief introduction of each tool will be made, as well as the most common advantages and disadvantages when working with this tool will be explained and finally for each tool it will be explained how it can be connected to SAP Hana to automate a process.

2. **PENTAHO**

Introduction to Pentaho

Pentaho is a solution-oriented, process-centric Business Intelligence (BI) platform that includes the components required to implement process-based solutions such as data mining, ETL, reporting etc.

It helps to improve analysis and decision-making capabilities, since it offers solutions that are mainly composed of an infrastructure of analysis and reporting tools integrated with a business process workflow engine.

ADVANTAGES AND DISADVANTAGES

When talking about the **advantages** of this tool we can see that:

- It's open source, so its not necessary to purchase a license for its use.
- It's multi-platform
- It uses standard technologies such as Java, XML or JavaScript
- It has basic development environments in PDI as well as in BI server and other tools.
- Easy to install and configure
- It is a flexible solution that allows the creation of new functionalities or modules that adapt to the needs of the organization for example by the installation of plugins to extend its functionality.
- It has a powerful OLAP query engine, such as Mondrian, which is also Java-based and allows us to perform powerful queries in a tiny amount of time.
- Focused on Big Data and non-relational databases.
- It is compatible with a multitude of data sources and Big Data systems, such as Hadoop and Apache Spark, which makes it very easy to connect it with different sources without having to adapt or go through third-party tools.



When talking about the **disadvantages** of this tool we can see that:

- It is a suite of tools that are designed by different developers, so we are going to have different menus, of different shapes and different visualizations, which can complicate a bit the learning in new Pentaho tools.
- There is not much reliable documentation, so many times we must rely on other users who have had the same error or the same need.
- There is only a basic user manual that does not detail many concepts

CONNECT ETL PROCESSES TO SAP HANA

By default, Pentaho has no connector to SAP databases. Therefore, to connect to these databases it is necessary to use the "Generic Database" connector, which connects via JDBC.

This connection requires a jdbc driver called "ngdbc" that must be inserted in the "lb" folder of the PDI .

It can be obtained from the SAP HANA STUDIO (Eclipse customized to work with SAP) and can also be downloaded from the following URL: <u>https://jar-download.com/?search_box=ngdbc</u>

The fields to be filled in this type of connections are:

- **Custom URL** : In this field you must specify the IP address of the SAP server and the port through which to connect to the corresponding tenant (see documentation or ask the SAP provider for the values). The formula to specify these aspects is as follows: jdbc:sap://<IP address>:<Port>.
- **Driver class:** In this field the following string must be inserted: com.sap.db.jdbc.Driver
- Username: The username authorized to access the database.
- **Password:** The password associated with the user.



••••••	•
Password:	
SAPHANADB	\$
Username:	
Authentication	
com.sap.db.jdbc.Driver	•
Custom driver class name:	
jdbc:sap://52.137.7.147:30241	•
Custom connection URL:	
Generic database	*
Dialect:	
Settings	

IT NOVUM CONECTORS FOR PDI

On the other hand, to connect to SAP from PDI, just know that IT Novum has developed a series of plugins to connect to internal data sources of some SAP solutions, only for SAP ERP, SAP BW and SAP R3/ERP.

All connectors require the same parameters to establish the connection:

- Host Name: IP address or URL of the SAP server.
- **System Number**: Two-digit number defining the instance number (usually 80 in systems). To know it you can use the "GET_SYSTEM_NUMBER" function.
- **SAP Client**: The client ID is a three-digit number. It can be obtained with the transaction "SC4" in the SAP GUI
- Username: User with permission to access the SAP system.
- **Password:** The password associated with the user.

Settings	
Host Name:	
10.0.0.148	•
System Number:	
00	•
SAP Client:	
001	•
Language:	
EN	\$
Username:	
DEVELOPER	\$
Password:	
•••••	\$



In this case for making the installation refer to the IT-Novum documentation.

To make the connection work, PDI must be installed on the virtual machine that had the SAP client and is linked to the server on which the system is running.

DATA EXTRACTION FROM AN SAP ERP OR SAP R3 SYSTEM – IT NOVUM

These connections are used to connect to ERP systems or to R3 systems. The connectors are used by the "SAP ERP Table Input" component, which is responsible for extracting data from the SAP ERP tables.

To use this tool, a connection to one of the two systems must be specified. The connection is made through the different connectors provided by IT Novum in its plugins.

Once the connection is established, the table from which the data is to be extracted must be identified. This identification can be done by entering the complete name of the table or by leaving the field blank and selecting the magnifying glass to the right of the field; in this case a pop-up window will appear with all the tables that you want to extract the data from.

Step name	SAP ERP Table Input					
Connection	SAP ERP	~	Edit	New	Wiza	rd
Table					•	٩

When the table is selected, in the lower part of the window will appear all the fields that the table has, the table with its name, the type of data, both for SAP and its transformation into PDI, its length and its field description. The "Convert" column specifies whether the data has been passed to the Java field type (specified in the "PDI-Data Type" field). This conversion can be selected with the different buttons below the fields.

These fields can be removed from the window in order not to include them in the ETL process. To do this, click on the field and press the delete button.



Table //BOBF/D_PR_NAME									
Gene	eral Advanced Filter	Server Mode							
#	Name	SAP-DataType	PDI-DataType	Length	Convert	Description			
1	MANDT	C	String	3	YES	Client			
2	DB_KEY	Х	String	16	YES	NodelD			
3	PARENT_KEY	х	String	16	YES	NodelD			
4	LANGUAGE_CODE	С	String	1	YES	Language Key			
5	NAME	С	String	255	YES	EPM: Text field for names and descriptions			
6	DESCRIPTION	С	String	255	YES	EPM: Text field for names and descriptions			

At the bottom of the window there are three buttons. The "Convert Yes" and "Convert No" buttons are the ones that define the value of the "Convert" column. Finally, the get fields button is used to retrieve the deleted fields.

POSITIVE AND NEGATIVE ASPECTS REGARDING THE SAP CONECTION

Pentaho is a powerful tool when offering ETL solutions that allow to integrate information in Data Lakes, Data Warehouses as well as to create applications and analytical projects. In addition, it also allows to extract all the necessary information of the SAP systems to be able to establish these connections.

However, since it does not have a specific connector for the SAP environment, it can be difficult to find documentation on errors and more specific information on how to use this component in the environment itself.

Although there is the possibility of installing a plugin in PDI to connect specifically to SAP solutions the process is somewhat complex so it is not worth this solution considering that there are other tools that can solve the automation of the process in a simpler way.



3. TALEND

Introduction to Talend

Talend Open Studio (TOS) is a suite that provides a very complex, varied and complete set of tools to perform data integration offered in an open-source version. It is also a Data Integration (ETL) platform that manages and implements new digital processes in companies, which will give you a competitive advantage.

It allows processes such as advanced analytics and decision making, it is a tool that gives Artificial Intelligence to existing or new processes, as well as implementation and improvement of ecommerce processes, digital marketing processes focused on target audiences, etc.

ADVANTAGES AND DISADVANTAGES

When talking about the **advantages** of this tool we can see that:

- It is an all-in-one, what allows us to reduce the number of tools and therefore additional configurations
- It has a relatively long-lived and well-structured documentation and community
- The use of generic components allows Talend to connect to virtually any platform
- Talend provides its own framework for developing custom components.

When talking about the **disadvantages** of this tool we can see that:

- The installation of the tool can be complex
- When first using the tool you may notice a steep learning curve due to its wide range of functionalities
- When using the tool, the consumption of machine resources can be quite high.
- The tool does not support the .ods format

CONNECT ETL PROCESSES TO SAP HANA

Talend does offer a specific connector to the SAP Hana database. This connector is also based on the JDBC connection. This connection requires the same jdbc driver that is used in the PDI connection. This driver must be inserted into the folder "configuration\m2\repository\org\talend\libraries\ngdbc" folder inside the Talend directory. The Driver can be obtained from the SAP Hana Studio, it can also be downloaded at https://jar-download.com/?search_box=ngdbc

In the Community version of Talend, this connection can be made with the component "tDBInput"



The fields to be filled in this type of connections are:

- **DB Type**: "SAPHana".
- Login: Authorized username to access the database.
- **Password:** The password associated with the user
- **Server:** IP address or URL of the SAP server.
- **Port:** Port through which to connect to the corresponding tenant.
- Schema: Database to connect to.

	~
jdbc:sap://52.137.7.147:30241?	^
SAPHANADB	
•••••	
52.137.7.147	
30241	
SAP_HANA_DEMO	
	~
Test connection	v
Exportar como contexto Revertir Contexto	
	jdbc:sap://52.137.7.147:30241? SAPHANADB S2.137.7.147 30241 SAP_HANA_DEMO Test connection Test connection

In the paid version of Talend there is a component that performs the same function without having to fill out the DB Type parameter. This component is called "tSAPHanaInput"

This component offers not only the ability to connect to SAP Hana databases but also the functionality to connect to different data structures used by SAP in its SAP Hana databases. With this component it will be possible to connect to any type of SAP system and extract data from its tables, its InfoProviders or its Bapis (SAP standard interfaces)

To be able to establish the talend connections with SAP systems, it is necessary to have in the directory "Talend-Studio v7.2.1\ configuration\ m2\ repository \org\talend\libraries\sapjco3\6.0.0" the driver JAR "sapjco3.jar" and in the "C:\Windows\System32" directory the "sapjco3.dll" file.

Afterwards, open the Talend tool, and in the upper part of the tool, drop down the "Window" option, then select the "Show view" option.



Then a new window will appear with several folders from which you should select the "Talend" folder and click on the "Modules" option.



With the modules window open, select the icon located in the upper right part of the modules window. After clicking on the icon, a new window will open where the path to the "sapjco.jar" file should be displayed.

뤎 Modu	R Modules × 🕹 🖓 ₹ ⊡ €								
Status	Context	Module	Description	Required	-				
🗹 Insta	plugin:org.talend.libraries.apache.axis2	activation-1.1.jar		\checkmark	Ξ				
🗹 Insta	tSAPBWInput	activation-1.1.jar	Required for using this component.	\checkmark					
🗹 Insta	tMicrosoftCrmInput	activation-1.1.jar	Required for using this component.						
🗹 Insta	tMicrosoftCrmOutput	activation-1.1.jar	Required for using this component.						
🗹 Insta	plugin:org.talend.libraries.apache.axis2	activation.jar		\checkmark					
🗹 Insta	tSugarCRMInput	activation.jar	Required for using this component.	\checkmark					
🗹 Insta	tWebService	activation.jar	Required for using this component.	\checkmark					
🗹 Insta	tMDMConnection	activation.jar	Required for using this component.	\checkmark	-				

On the other hand, with Talend version 7.2 it is not installed by default the SAP connector that is used to extract data from InfoProviders generated by SAP BW. So, without installing this connector you will be able to see that there are InfoProviders created but you will not be able to extract the information stored in those InfoProviders.

The installation process can be quite tedious, as you must work with both the SAP client and the server where it is installed. To install this connector, you must go to the path "\TalendStudio7.2.1\plugins\ org.talend.libraries.sap_7.2.1.20190419_0324 \resources"

in it there are several compressed files with different components that can be installed to extend Talend's functionality, and, in addition, there is a text file called "readme"



where the process to be followed to install these features is explained. The file to be used is the "TIDK900022" file which contains two other files: the

"K9000220.TID" and "R900022.TID". These two files are the ones you must put in the server where the system is installed.

To access the system, you must connect via SSH, and once inside you must navigate until you reach the directory "trans" to the path "/usr/sap/trans". Inside this directory there are several folders created, but the only ones that are going to be modified are the "data" and "cofiles" folders, since the "cofies" folder is where the "K900022.TID" file must be inserted and in the "data" folder the "R900022.TID " file must be inserted. Once this operation has been carried out, you will not work with the server again if there are no problems.

When working with the server, you can check with the transaction "AL 11" that the two files have been correctly entered into the system.



In the transaction we will see all the SAP directories and in between of them we will see the "DIR_TRANS" one. If you click in it you can see what it contains



Image: Second	了 III - 任 · 编								
Name of Directory Parameter DIR_ATRA DIR_BINARY DIR_CCMS DIR_CT_LOGGING	Directory								
DIR_ATRA DIR_BINARY DIR_CCMS DIR_CT_LOGGING	Directory								
DIR_BINARY DIR_CCMS DIR_CT_LOGGING	/usr/sap/NPL/D00/data	Direc	tory: /usi	r/sap/tra	ans/c	ofiles			
DIR_CCMS	/usr/sap/NPL/D00/exe	D. 07				-		AA - 5	
DIR CT LOGGING	/usr/sap/ccms	😏 🖳	₩ 🦻 A 🤋	5 TO X 🛙 🖸				-w 🚞	
	/usr/sap/NPL/SYS/global								
DIR_CT_RUN	/usr/sap/NPL/SYS/exe/uc/linuxx86_64	Usable Vie	ewed Changed	Length Owr	er Last	change	Lastc	hange File Na	me
DIR_DATA	/usr/sap/NPL/D00/data	6		4096 npla	dm 17.0	6.2020	11:17	/:30 .	
DIR_DBMS	/usr/sap/NPL/SYS/SAPDB			4096 nola	dm 03.0	6.2020	17:13	3:29	
DIR_EXECUTABLE	/usr/sap/NPL/D00/exe	Х		580 npla	dm 19.0	6.2020	11:14	:05 K9000	22.TID
DIR_EXE_ROOT	/usr/sap/NPL/SYS/exe								
DIR_GEN	/usr/sap/NPL/SYS/exe/dbg								
DIR_GEN_ROOT	/usr/sap/NPL/SYS/gen								
DIR_GLOBAL	/usr/sap/NPL/SYS/global								
DIR_GRAPH_EXE	/usr/sap/NPL/D00/exe				_ \				
DIR_GRAPH_LIB	/usr/sap/NPL/D00/exe								
DIR_HOME	/usr/sap/NPL/D00/work								
DIR_INSTALL	/usr/sap/NPL/SYS								
DIR_INSTANCE	/usr/sap/NPL/D00	Usable	Viewed Chang	ged Length	Owner	Lastchar	nge	Lastchange	File Na
DIR_LIBRARY	/usr/sap/NPL/D00/exe			4096	npladm	03.06.	020	17:13:29	
DIR_LOGGING	/usr/sap/NPL/D00/log			4096	root	03.06.2	0.20	17:13:28	
DIR_MEMORY_INSPECTOR	/usr/sap/NPL/D00/data			4096	nnladm	03.06.2	020	17-13-29	FPS
DIR_PAGING	/usr/sap/NPL/D00/data			4006	nobdm	02.06.2	020	17:12:20	action
DIR_PERF	/usr/sap/tmp			4090	npiaum	03.00.2	020	17:13:28	actiog
DIR_PROFILE	/usr/sap/NPL/SYS/profile			4096	npladm	19.06.2	020	10:13:42	bin
DIR_PUT	/usr/sap/put			4096	npladm	19.06.2	020	10.46:49	buffer
DIR_REORG	/usr/sap/NPL/D00/data			4096	npladm	17.06.2	020	11:17:30	cofiles
	/usr/sap/NPL/D00/exe			4096	npladm	17.06.2	020	11:18:41	data
DIR_SAPHOSTAGENT	/usr/sap/nostctri			4096	noladm	03.06.2	020	17.13.28	etc
DIR_SAPUSERS	·			4006	nobdm	10.06.2	020	10.56.5	laa
DIR_SETUPS	/usr/sap/NPL/SYS/profile			4090	npiaum	19.00.2	020	10:56:1	log
	/usr/sap/NPL/D00/data			4096	npladm	16.06.2	020	14:06.17	sapnar
DIR_SOURCE	/usr/sdp/NPL/SYS/SrC			4096	npladm	03.06.2	020	17 13:28	storag
DID TRANC	lucionitran			4096	npladm	19.06.2	020	2:02:27	tmp

Once you have verified that the two files are in their corresponding directories, you must switch to the "STMS" transaction, where the transaction will be activated with the two files that have been inserted into the system.

BIG DATA – BUSINESS INTELLIGENCE





🖉 STMS 🔹 🔍 🖷 ା 🗞 🚷 ଜୋଳ ଲିଲି । ସିସିରି ସି । 🜄 🗖 । ଡେ 🎙
Transport Management System
<mark>∍</mark> \$≿⊘ &\$ \$ 4 ¥ ∎
System NPL NPL Local Transp. Domain DOMAIN_NPL Transport domain NPL
*
TRANSPORT MARAGEMENT SYSTEM

In the new window, select the physical system you are working with by double clicking on the system. You can see the system you are working with from the initial window.

Import Overview: Domain DOMAIN_NPL									
🔁 🆆 🗞 🛫 🊟 🎵 🚰 📰 🔝									
📑 Number	r of import queues: 2								
Queue	Description								
CL5 🕺 CL5 NPL 🛐 NPL Local									
	1								

Within the selected system the queue with all transport activation requests is displayed. To create a new transport request, go to the upper part of the window, where you will find the extras option, select the "Other requests" drop-down menu and in the drop-down menu select the "Add" section.



After selecting this option, a pop-up window will appear where the name of the transaction to be added to the queue must be entered. The name of the transaction will be the same as the name of the compressed file you worked with at the beginning of the process: "TIDK900022



C Add Transport Re	equest to Impor	t Queue	×
Transp. Request Import Queue	TIDK9000	NPL Local	
Import Again		_	
			2

To activate it you must press the f9 key on your keyboard and then press the "ctrl" and "F11" keys at the same time. This request triggers a new pop-up window where you must specify some aspects of the request. In this window it is important to correctly select the user number that is being used in the system.

Requests for NPL: 0 / 1									
Number	Request	T	QM	RC	I	UMC	Owner	Project	Short Text
1	TIDK900022	ĸ	100	٠	k	1	DRASSEBAUM		

ransport Request	TID#900022
arget System	NPL NPL Local
arget Clent	Tarp.Clent=Source Clent
Date Execut	ion Options
Chut Data	
O famou di sha	
• Immediate	
OAt Start Time	
Planned Start	19.06.2020 (11:02:38)
No Start After	
OAfter Event	
Event	
Parameters	

Once the Talend connectors are integrated in SAP you can extract data from the system

In Talend you can create a connection to an SAP solution in several ways, but the easiest is from the Metadata tool in the "Repository" window (on the left side of the screen). In the Metadata tool there is an option called "SAP Connections", which is the tool that should be used to connect to SAP systems.

Right-click on this tool and you will see a small drop-down menu in which one of the options is "Create SAP connection". Selecting this option will pop up a new window with the parameters to be filled in to make the connection.





The window where the connection parameters must be filled in has three steps:

- The first step is where you name the connection and where you can put a small description of the connection.
- The second step is where you fill in the parameters needed to establish a connection to the system:
 - **Client**: This is the ID of the client and is a three-digit number. It can be obtained with the transaction "SCC4" in the SAP GUI.
 - Host Name: IP address or URL of the SAP server.
 - **User:** User with permission to access the SAP system.
 - **Password:** The password associated with the user.
 - System Number: Two-digit number that defines the instance number. To know it you can use the GET_SYSTEM_NUMBER function
 - Language: The language in which the system is usually English (EN) or German (DE).
 - Additional Properties: It is recommended to add the property "api.use_z_talend_read_ table" with its value "true" so that there are no problems with the maximum data length supported by Talend when reading tables.

0	SAP Connection
Jpdate SAP Conn	ection - Step 2/2
Update SAP connectio	n parameters
Client	001
Host	10.0.0.57
User	DEVELOPER
Password	*****
System Number	00
Language	EN
Additional Properties	Property Name Value
Help	< Back Next > Finish Cancel



• The third step is optional and is to connect to the SAP Hana database linked to the system.

SAP Connection	-		x
Update SAP Connection - Step 3/3		N	
Db Host must be specified		M	
SAP Hana Database connection parameters			
Db Host			
Db Port			
Db Schema			
Db Username			
Db Password			
Check Export as context Revert Context			
Help < Back Next > Finish	С	ancel	

After creating the connection, if we display the option in the "SAP Connections" tool located at Metadata, you will be able to visualize the connection created. This connection has a drop-down appearance and if it is displayed, all the types of data structures that the SAP system has from which data can be extracted are shown.

⊿ 🖅 SAP 0.1
SAP BI Content Extractor
SAP Tables
🗀 SAP Bapi
SAP iDocs
SAP ADSO
SAP DataSource
SAP DSO
SAP InfoCube
SAP InfoObject

To extract the data from these structures, right-click on the connection (in the case of the example on "SAP 0.1") and select the structure from which you want to extract the data.



The option we are going to work with in this document is the "Retrieve SAP Table", whose functionality is to extract data from a table in the SAP system.

SAP	Edit SAP connection
==	Retrieve SAP Table
==	Retrieve Business Content Extractor
SAP	Retrieve Bapi
==	Retrieve SAP BW metadata
==	Create SAP IDoc
	Import Data Mapper IDoc Structures
\$	Impact Analysis
\$	Detect dependencies
×	Delete
	Сору
6	Duplicate
Q	Export items

RETRIEVE SAP TABLE

With this tool you can extract data from the internal tables that the SAP solution has on the server.

When you select this tool a new window pops up with a Wizard with several steps in which you will specify several parameters to find the data you want to extract.

In the first step you identify the tables from which you want to extract data. The
window where you select these tables has in the upper part two spaces where
you can insert the table name or its description. If neither the name nor the
description is known, you can fill it with an asterisk and all the tables in the
system will be selected. To perform the search, you must click on the search
button to the right of these fields. Below the two fields there is a box in which all
the tables corresponding to the search will appear. The name and description of
these tables are displayed and if selected, the number of columns will also be
displayed.



• The second step shows the fields contained in the selected tables and a small "Preview" of their data.

Search Table					
Name		Descriptio	n _*		Sear
Name	Description	Туре	Column Number	Creation Status	
CDPOS	Change document items				
DOKTL	Documentation - text li				
CVT3	Cluster VERI-Table				
EDID2	IDoc Data Record from				
EDID4	IDoc Data Records from				
EDIDD_OLD	IDoc Data Record				
DSYGH	DSYS: Parent-Child Tabl				
DSYGI	DSYS: Grouping Admini				
DSYGL	DSYS Parent-child Table	TRANSPARENT	7	Success	
DSYOL	SAPfind - DSYS, Module				
DSYOT	SAPfind DSYS: Chapter				
PCDPOS	Change document items				
SFHOT	SAPfind: Active Structures				
TERMC1	Terminology entries, de				
TERMC2	Glossary terms, definitio				
TERMC3	Terminology entries, cr				

Once the selection of tables is finished, deploying the connection to the system, the section of SAP Tables will also be displayed, and you will be able to see all the tables in the system that have been selected.



To work with them in the ETL process, all you must do is drag the table to the job design you are working on. When the table has been dropped on the job design, a new window will appear showing the tools that can be used to work with the table. Select the option "tSAPTableInput".





POSITIVE AND NEGATIVE ASPECTS REGARDING THE SAP CONECTION

Talend, therefore, is a complete low-code platform that unifies and manages the organization's data lifecycle, backed by professional services and ecosystem partners and built for cloud, multi-cloud and hybrid environments.

In this case, unlike Pentaho, Talend does have a specific connector for the SAP environment, both in its community version and in its paid version, which could be a more successful solution when looking for a tool that allows us to connect to SAP to automate a process.

In most cases the community version is enough, it all depends on the SAP solution to which you want to connect and the complexity of the process you want to perform. On the other hand, from the paid version of Talend, the steps to establish the connection itself are somewhat more complex to accomplish so it can be tedious to use this tool if the process to be performed is something simple that could be done with another tool.

In summary, considering that the installation of the tool is complex, that it is even more complicated to use it due to its great variety of functionalities and that the consumption of machine resources can be quite high, it is not worth using this tool for simple processes that could be executed with another tool.



4. **POWER AUTOMATE**

Introduction to Power Automate

Power Automate is an enterprise system from Microsoft, it is part of Microsoft Power Platform and allows us to integrate and synchronize all our data analysis and applications in an automated way, with the objective of increasing productivity and business efficiency.

In this way, we can simplify repetitive tasks and business processes so that we can focus our attention on what is most needed.

ADVANTAGES AND DISADVANTAGES

When talking about the **advantages** of this tool we can see that:

- There are many templates available
- Is extremely easy to use, even for non-technical users
- Easily integrates with other aplications thanks to the plenty of built-in conectors that it gathers and the possibility of creating custom conectors
- Power Automate helps users to automatize and prioritize their tasks
- It allows users to easily share and access data by connecting your favorite services together and creating a flow between them.

When talking about the **disadvantages** of this tool we can see that:

- When creating complex workflows, it affects the performance of the program considerably
- Maximum flow frequency for a free plan is 15 minutes as compared to app-based plan which is 1 minute. Flow frequency is the time it takes for the flow to trigger.
- Power Automate supports only sequential workflows.
- You can only save workflow instances for 30 days which might not be adequate for business-critical processes.
- You need a premium license to access to the SAP conector and use Power Automate Desktop

CONNECT ETL PROCESSES TO SAP HANA

- Power Automate Cloud service:
 - To make the connection to SAP HANA it is necessary to install an onpremises gateway as well as create a user, with certain permissions and



a password through which we will connect to the SAP system from Power Automate.

• Then we will select the type of flow that we want for our process to be. In this case we are going to select an instant flow so when a user for examples clicks a button the flow will run.





 Once we create the flow we select the next step option, and we will look for the SAP ERP conector. The SAP ERP connector for Power Automate allows us to connect our workflows to an SAP ERP system (either SAP ECC or SAP S/4HANA). Once we chose the conector we can see that there are plenty of actions available. We need to take into account that we will need a Power Automate license to use this type of connector.



SAP FRP	0
Search connectors and actions	
Triggers Actions	See mo
Call SAP function PREMIUM	
Call SAP function (V2) (preview) PREMIUM SAP ERP	
Create stateful session (preview) PREMIUM SAP ERP	
Read SAP table with parsing (preview) PREMIUM	
Run Diagnostics PREMIUM SAP ERP	
Don't see what you need?	
Help us decide which connectors and triggers to add next with UserVoice	

- For testing the connection, we are going to select the one that's called "Call SAP function."
- Before filling the fields for the proper action, we are going to add the SAP connection to the Power Automate flow. For this connection there are different fields:
 - Authentication Type: SAP Authentication
 - Data Gateway: The one-premise gateway that we have created previously
 - SAP Username
 - SAP Password

Manually trigger a flow		0	
	+		
SAP ERP		()	
* Authentication Type	SAP Authentication		\mathbf{v}
0			
* Data Gateway	PM1-on-prem-gw-on-DEMO-NEU-SAP-Us		\sim
0			
* SAP Username	pm1rfc		
0			
* SAP Password			
0			
	Creating Cancel		



- With this step we are checking our connectivity to the on-premises data gateway
- Once we establish our connection we can fill in the fields of the action that we have selected previously. The fields that we need to fill in are:
 - As Host: This is the Internal Ip address or internal host name
 - **Client:** This is the SAP Client ID
 - As System: This is the SAP System Number, that it is a number ranging from 00 to 99
 - SAP function name: You can select different functions for your flow depending on what you want to do with your data from SAP. When you select a specific function Power Automate automatically fetches all the relevant properties or fields that are in the system for that specific function.

-	¥	
Call SAP func	tion	• • •
* AS Host	10.7.0.5	
* Client	001	
* AS System Number	00	Show option
* SAP function name	BAPI_EPM_PRODUCT_GET_LIST	~
*Stateful Session	Yes	~
Use SNC	No	~
SNC library	Optional	
SNC SSO	Off	~
SNC My Name	Optional	
SNC Partner Name	Optional	
SNC Quality of Protection	Default	~
MAX_ROWS BAPIMAXROW		
HEADERDATA		
HEADERDATA		
HEADERDATA		

• Once we finish configuring our action we can test our flow with the test button and see not only if it works properly, but also the different data that it returns when calling the SAP function.





HEADERDATA	
"TYPE CODE", "DP"	-
"CATEGORY": "Notabooks"	
"NAME": "Notebook Basic 15".	
"DESCRIPTION": "Notebook Basic 15 with 2,80	GHz guad core, 15
"SUPPLIER_ID": "0100000000",	
"SUPPLIER_NAME": "SAP",	
<pre>"TAX_TARIF_CODE": 1,</pre>	,,
RETURN	
SELPARAMCATEGORIES	
0	
SELPARAMPRODUCTID	
Π	
SELPARAMSUPPLIERNAMES	
Π	
body	
1	-
"HEADERDATA": [_
{	
"PRODUCT_ID": "HT-1000",	
"TYPE_CODE": "PR",	
"LATEGORY": "Notebooks",	
"DESCONDITION", "Hat-hard Desig 15 with 2	00 Cits and care . 7

• Once we have set this action and verified that the flow returns data from the SAP environment we can continue to configure our data flow, depending on our business requirements.

In terms of licensing for using this solution we will need a Power Platform license (which in a lot of cases is already part of our Microsoft 365 license) and a license for using premium conectors because the SAP ERP conector is a premium conector. And on the SAP side we will need to have also the appropriate licenses for accessing de environment

POSITIVE AND NEGATIVE ASPECTS REGARDING THE SAP CONECTION

Power Automate is a tool to create simple workflows both in the cloud and in local environments with the Power Automate Desktop application. For our connection with SAP, we need only the cloud Power Automate tool.

As for the connection with SAP, if what you want is to create a simple workflow that connects to SAP and returns a file you could use this tool because it works for simple flows, and it is extremely easy to use even for non-technical users. However, it should be noted that a paid license is required to use it, since the connector that connects to SAP to take the data is a premium one.

On the other hand, it should also be noted that it requires an on-premises gateway (required version: December 2019 or higher) and an SAP .NET Connector SDK 3.0 from SAP. The access to the download requires a valid S-user. If you do not have an S-User account in the SAP environment it will be impossible to make the connection as it is



required as well as the on-premises getaway and therefore you will not be able to use the tool to create the workflow.

Also, there are some known issues and limitations for the SAP ERP conector:

- The connector supports only RFC's and BAPIs
- The connector does not support receiving messages from SAP Server
- Transactional RFCs (tRFCs) are not supported

If there is a lot of data to load into the file or the flow has a lot of steps, it can take much longer than expected and the flow can become an inefficient process.



5. **AZURE DATA FACTORY / SYNAPSE ANALYTICS**

Introduction to Azure Data Factory / Synapse Analytics

Azure Data Factory is Azure's cloud ETL service for horizontally scalable serverless data integration and transformation. It offers a code-free user interface that supports intuitive creation, monitoring and management from a single pane of glass.

Azure Synapse Analytics is a limitless analytics service that brings together data integration, enterprise data warehousing, and big data analytics. It gives you the freedom to query data on your terms, using either serverless or dedicated options—at scale.

ADVANTAGES AND DISADVANTAGES

- When talking about the **advantages** of this tool we can see that:
 - No-code data workflows: Configure Azure Data Factory to collect and integrate data from most data sources without having to write a single line of code.
 - Large collection of data collectors: Azure Data Factory currently offers nearly 100 prebuilt data connectors to import data from external sources. A large amount of Azure Data Factory's online data collectors can be set up instantly.
 - Built-in monitoring and alerting: Azure Data Factory offers built-in monitoring visualization. These native visibility features mean you can easily keep track of the status of data integration operations. On top of this, it helps the user be proactive about identifying and reacting to problems, such as a failed data transformation, that could disrupt workflows. You can also set up alerts to warn about such failed operations.
 - Consumption-based pricing: Unlike on-premises data integration tools, which typically require a large upfront investment, Azure Data Factory offers pay-as-you-go pricing.
 - Fully integrated with Azure Data services
 - Drag and drop UI
- When talking about the **disadvantages** of this tool we can see that:
 - Custom data collectors: While you can create data pipelines based on a variety of common sources without writing code in Azure Data Factory, you'll need to write custom code to configure nonstandard data sources.
 - Focus on Azure: Azure Data Factory supports some data sources hosted outside of Azure, but it's designed first and foremost for building



integration pipelines that connect to Azure or other Microsoft resource types.

 Long-term expense: While consumption-based pricing is attractive in some ways, its long-term total cost of ownership may be higher than that of on-premises options. If you plan to run data integration services for years, you may save money by hosting it on your own infrastructure.

CONNECT ETL PROCESSES TO SAP HANA

CONECTOR SAP HANA

The SAP HANA connector it's suitable for ingestion of data from SAP Hana databases. It works with SAP on premise and SAP on cloud, and it can connect also to information models like analytics and calculated views and has the capability of connecting to row column tables.

It runs on the Azure Data Factory self-hosted integration runtime, so it needs to be run on a virtual machine, specifically in a Windows Virtual Machine.

For conecting to SAP HANA, we need to install two things:

- SAP HANA ODBC driver installed in the virtual machine
- Azure Data Factory self-hosted integration runtime installed

The virtual machine needs to be in the same VirtualNet with SAP, and it must be able to connect to SAP using the HANA OBCD driver, so it must be in the same network within the SAP environment.

In summary, to be able to connect to S/4HANA we must follow these requirements:

- We need an Azure subscription
- We need access to our SAP HANA environment
- We need to have installed the SAP HANA client 2.0 which is the ODBC driver to connect to HANA (<u>https://learn.microsoft.com/en-us/power-bi/connect-data/desktop-sap-hana</u>) and you will need an Azure user ID on the SAP marketplace in order to download this

CONNECT AND COPY SAP HANA DATA WITH AZURE DATA FACTORY

• First thing that we need to do is go to Azure subscription and create a storage account and fill in all the fields



E 🗣 🖉 🎯 ? 🤅 Microsoft Azure , Search resources, services, and docs (G+/ All services & Search All Overview Featured Categories 0 ۲ 8 * 4> W All Virtual App Services Storage SQL Azure Azure Kubernetes Function App General Cosmos DE services Compute 0 \rightarrow $\langle \cdot \cdot \rangle$ 0 Storage accounts Networking Security Center Cost Virtua zure Activ All services Storage Manage Directory + Create O View Web Mobile Free training from Microsoft See all Containers Databases ۲ Analytics Core Cloud Services - Azure architecture and service guarantees Core Cloud Services - Manage services with the Azure portal Cloud Concepts - Principles of cloud computing Blockchain 9 units • 45 min 9 units • 1 hr 13 min 10 units • 1 hr 2 min Al + machine learning Azure provides a global network of sect datacenters you can deploy your service into. Learn about the physical architectu of Azure, how redundancy is provided, and what sort of service guarantees Explore the core concepts of cloud computing and how it can help your business. Tour the Azure portal features and services, and customize the portal. Internet of things Mixed reality Integration Start 🗹 Start 🖸 Start 🗹 Identity Carurit

- For the Basics page we will need to configure:
 - **Resource Group**: A container that holds related resources for an Azure solution. In our case we are going to create a new one.
 - Storage Account Name: A name for the storage account
 - **Location**: Where the SAP HANA instance is at, so that would be within the same data center
 - **Replication**: How you can replicate the data in your storage

Create storage account		
Basics Networking Advan	nced Tags Review + create	
Azure Storage is a Microsoft-manaç redundant. Azure Storage includes Tables. The cost of your storage acc Learn more about Azure storage ac	ed service providing cloud storage that is highly available, se Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Fi ount depends on the usage and the options you choose belo counts O ⁶	cure, durable, scalable, and lies, Azure Queues, and Azure w.
Project details		
Select the subscription to manage of your resources.	seployed resources and costs. Use resource groups like folder	s to organize and manage all
Subscription *	Microsoft Azure Sponsorship 2	~
Resource group *	(New) adf_lab	~
Resource group * Instance details The default deployment model is Ri the classic deployment model inste	(New) adf_lab Create new ssource Manager, which supports the latest Azure features. Yo d. Choose classic deployment model	su may choose to deploy usin
Resource group * Instance details The default deployment model is R the classic deployment model inste Storage account name * ①	[Pievi) adf_ab Create new source Manager, which supports the latest Azure features. Yo d. Choose classic deployment model adfsalesdata	su may choose to deploy usin
Resource group * Instance details The default deployment model is R the classic deployment model inste Storage account name *	(Mev) adf_ab Create new source Manager, which supports the latest Azure features. Yo ad. Choose classic deployment model adfsalesdata (IUS) East US 2	u may choose to deploy usin
Resource group * Instance details The default deployment model is R the classic deployment model inste Storage account name * Location * Performance	(New) adf_lab Create new ssource Manager, which supports the latest Azure features. Yo adfsalesdata (US) East US 2 Standard O Premium	nu may choose to deploy usin
Resource group * Instance details The default deployment model is R the classic deployment model inste Storage account name * Location * Performance Account kind	(New) adf_lab Create new ssource Manager, which supports the latest Azure features. Ye adfsalesdata (US) East US 2 € Standard Premium StorageV2 (general purpose V2)	u may choose to deploy usin
Resource group * Instance details Instance details The default deployment model in Re Storage account name * Location * Performance Account kind Replication	(New) adf_lab Create new ssource Manager, which supports the latest Azure features. Yo adfsalesdata (US) East US 2 ③ Standard ○ Premium StorageV2 (general purpose V2) Locally-redundant storage (LRS)	su may choose to deploy usin
Resource group * Instance details Instance details The default deployment model is R the classic deployment model inste Storage account name * Location * Performance Account kind Replication Access tier (default)	[Ptevi) adf_lab Create new source Manager, which supports the latest Azure features. Yo adfsalesdata [U5) East U5 2 ③ Standard ○ Premium StorageV2 (general purpose v2) Locally-redundant storage (LRS) ○ Cool ④ Hot	u may choose to deploy usin

• For the Networking page we will need to configure:



• Connectivity method: Where you can connect your data

All services > Create storage acco	unt
Create storage account	
Basics Networking Adv	anced Tags Review + create
Network connectivity	
You can connect to your storage a private endpoint.	account either publicly, via public IP addresses or service endpoints, or privately, using a
Connectivity method *	Public endpoint (all networks)
	 Public endpoint (selected networks)
	O Private endpoint
	O All networks will be able to access this storage account. Learn more about connectivity methods □ [*]

- For the Advanced page we will need to configure:
 - Security
 - Large files shares
 - Blob soft delete
 - Hierarchical namespace

= Microsoft Azure	Search resources, services, and docs (G+/)
All services > Create storage account	
Create storage account	
Basics Networking Advanced	Tags Review + create
Security	
Secure transfer required ③	O Disabled 💽 Enabled
Azure Files	
Large file shares ①	Disabled Enabled
Data protection	
Blob soft delete ①	Disabled Enabled
Versioning ①	Disabled Enabled
	 The current combination of subscription, storage account kind, performance, replication and location does not support versioning.
Data Lake Storage Gen2	
Hierarchical namespace ①	Disabled Enabled
NFS v3 ①	Disabled Enabled
	O Sign up is currently required to utilize the NFS v3 feature on a per-subscription basis. Sign up for NFS v3 C ^a

• For the Tag page we will not need to configure anything



Il services > Create stor	age account	
Create storage ac	count	
Basics Networking	Advanced Tags Review + cre	nate
Tags are name/value pair multiple resources and re	s that enable you to categorize resources a source groups. Learn more about tags 🗆	and view consolidated billing by applying the same tag to
Note that if you create ta	gs and then change resource settings on o	other tabs, your tags will be automatically updated.
Name 🛈	Value 🛈	Resource

 For finishing configurating the storage account we will go to the Review and create page where we can see a resume of the storage account that we have created and once we click the button create it will be automatically created

Create storage account							
Validation passed							
Basics Networking	Advanced	Tags Review + create					
Basics							
Subscription		Microsoft Azure Sponsorship 2					
Resource group		(New) adf_lab					
Location		East US 2					
Storage account name		adfsalesdata					
Deployment model		Resource manager					
Account kind		StorageV2 (general purpose v2)					
Replication		Locally-redundant storage (LRS)					
Performance		Standard					
Access tier (default)		Hot					
Networking							
Connectivity method		Public endpoint (all networks)					
Advanced							
Secure transfer required		Enabled					
Large file shares		Disabled					
Blob soft delete		Disabled					
Blob change feed		Disabled					
Versioning		Disabled					
Hierarchical namespace		Disabled					
NFS v3	13	Disabled					
Create	< Previ	ous Next > Downloa					

• Now we need to create the virtual machine. We will go back to the "All services" page in Azure and select the virtual machine option.





 Dearus resources, 	services, and docs (G	i+/)					E G
Featured							
P			SQL	97	3	000	4>
Virt ^(h) machines	App Services	Storage accounts	SQL databases	Azure Database for	Azure Cosmos DB	Kubernetes services	Function App
{·· >	٨				0	0	\rightarrow
Virtual networks	Azure Active Directory	Resource groups	Monitor	Advisor	Security Center	Cost Management	All services
	Featured	Featured Virið machines Virið App Services Azure Active Directory	Featured With App Services Stronge accounts Virtual Virtual Active Active Groups Stronge	Featured Virt ^b machines Virt ^b App Services Storage accounts Storage Stor	Featured Virt ^b machines Virt ^b Wirk ^b App Services Storage accounts Storage Stor	Featured Virtib Wirb App Services App Services App Services App Services Storage accounts databases Virtual V	Featured With Age Services Storage SQL Azure Azure Azure Services Storage SQL Azure Database for Courses DB Services Services Services Courses DB Services Services Courses Courses DB Services Services Courses Courses DB Services Services Courses DB Services Services Courses DB Services Courses DB Services Services Courses DB Services Services Courses DB Services

- For the Basics page we will need to configure:
 - **Resource Group**: A container that holds related resources for an Azure solution.
 - Virtual Machine Name: A name for the storage account
 - **Region**: Where the SAP HANA instance is at, so that would be within the same data center
 - **Image**: The type of virtual machine that we want.
 - Size: The size of our virtual machine
 - Username: The username for log in the virtual machine
 - **Password**: The password for log in the virtual machine
 - **Selected inbound ports**: Configure the ports that we want to select for our virtual machine

Create a virtual machine		
Instance details		
Virtual machine name * 🕕	adfvm	~
Region * 🕕	(US) East US 2	×
Availability options 🛈	No infrastructure redundancy required	~
Image * 🕕	Windows Server 2016 Datacenter	V
	Browse all public and private images	
Azure Spot instance ①	🔿 Yes 💿 No	
Size * ①	Standard D4 v3	
	4 vcpus, 16 GiB memory (\$274.48/month)	
	Change size	
Administrator account		
Username * 🛈	adfadmin	~
Password * 🕕		~
Confirm password * ①		~
Inbound port rules		
Select which virtual machine netwo network access on the Networking	ork ports are accessible from the public internet. You can specify more limited or gra tab.	nular
Public inbound ports * 🛈	None Allow selected ports	
Calact inhound ports *	000 (2200)	

- For the Disks page we will need to configure:
 - OS disk type
 - Encryption type



All services	> Virtual machines > C	reate a virtual machine		
Create a	virtual machine			
Basics	Disks Networking	Management A	dvanced Tags F	eview + create
Azure VMs The size of	have one operating syste the VM determines the t	em disk and a temporary ype of storage you can u	disk for short-term sto ise and the number of d	rage. You can attach additional data disks ata disks allowed. Learn more
Disk optio	ns			
OS disk typ	e* 🛈	Standard SSD		
Encryption	type *	(Default) Encry	ption at-rest with a plat	orm-managed key
Enable Ultr	a Disk compatibility ①	🔿 Yes 🛞 N	0	
		Ultra Disk comp	atibility is not available	for this VM size and location.
Data disks				
You can add temporary o	d and configure addition disk.	al data disks for your virt	tual machine or attach e	xisting disks. This VM also comes with a
	Name	Size (GiB)	Disk type	Host caching
LUN		tach an existing disk		
LUN Create and	attach a new disk At			

- For the Networking page we will need to configure:
 - Virtual Network
 - Subnet
 - Public IP
 - NIC network security
 - Select inbound ports

Dusics	Disks	Networking	Management	Advanced	Tags	Review + create		
Define ne ports, inb Learn mo	twork con ound and ore	nectivity for your outbound connec	virtual machine by tivity with security	configuring net group rules, or	work inte place bei	rface card (NIC) settings. You can co ind an existing load balancing solu	ontrol tion.	
Network	interface	i i						
When cre	ating a virt	tual machine, a ne	twork interface wi	II be created for	you.			
Virtual ne	twork *	⁰	(new) adf	lab-vnet			~	
			Create new					
Subnet *	0		(new) defa	ult (10.1.9.0/24)			~	
Public IP	0		(new) adfv	(new) adfym-ip				
			Create new					
NIC netw	ork securi	ty group 💿	O None	● Basic 〇	Advance	d		
Public int	ound por	ts * 🛈	O None	 Allow select 	ted ports			
Select inb	ound por	ts *	HTTP (80), HTTPS (443), RDP (3389)					
			A This record to create	will allow all IP a nmended for tes eate rules to limit	ddresses ing. Use inbound	to access your virtual machine. This the Advanced controls in the Networ traffic to known IP addresses.	is only king tab	
	ed networ	king ①	• On () off				
Accelerat								

- For the Management page we will need to configure:
 - Boot diagnosis
 - OS guest diagnosis



- Diagnosis storage account
- System assigned managed identity

Create a virtual machi	ne			
Basics Disks Network	ing Management	Advanced	Tags	Review + create
Configure monitoring and mar	agement options for your	VM.		
Azure Security Center				
Azure Security Center provides Learn more	unified security manager	nent and advan	ced threa	t protection across hybrid cloud workloads.
Your subscription is prote	ted by Azure Security Cer	nter standard pl	an.	
Monitoring				
Boot diagnostics 🛈	⊙ On ()) Off		
OS guest diagnostics 🕕	() On ()	Off		
Diagnostics storage account *	() adfsalesdat	a		~
	Create new			
Identity				
System assigned managed ide	ntity 💿 🔿 On 💽	Off		
Azure Active Directory				
Login with AAD credentials (P	review) () () On ()	Off		

• For the Advanced and Tag page we are going to leave the default settings and then we are going to go to the Review + Create page for finally creating the virtual machine:

Microsoft Azure	℘ Search resources, services, and docs (G+/)
All services > Virtual machines > 0	Treate a virtual machine
Create a virtual machine	
Validation passed	
Basics Disks Networking	Management Advanced Tags Review + create
PRODUCT DETAILS	
Standard D4 v3	Subscription credits apply ③
by Microsoft	0.3760 USD/hr
Terms of use Privacy policy	Pricing for other VM sizes
TERMS	
Information with the provider(s) of provide rights for third-party offeri	of public and (1) agree that individue had barre in y counts, and the diage and unuskictional the offering(s) of support, billing and other transactional activities, Microsoft does not sgs. See the Azure Marketplace Terms for additional details.
Basics	
Subscription	Microsoft Azure Sponsorship 2
Resource group	adf_lab
Virtual machine name	adfvm
Region	East US 2
Availability options	No infrastructure redundancy required
Username	adfadmin



• Now that we have created the storage account and the virtual machine we want to create a container for the storage account. For doing that we are going to click the button "Go to resource"



• Then we are going to select the name of the resource group that we have created



 The next step it to look for the Storage account resource in our resource group



			еф 🖓 🧐 🤤 зарнанамтеореморитьсок	: 🔍
All services > CreateVm-MicrosoftWindows	(Server, WindowsServer-201-20200509210531 Overview >	adtvm > adt_lab	Å	? ×
	+ Add III Edit columns 📋 Delete resource gro	up \bigcirc Refresh \rightarrow Move $\stackrel{1}{=}$ Export to CSV $ $ \oslash Assign tags \square Del	lete 🛓 Export template 💙 Feedback	
(e) Overview	Subscription (change) : Microsoft Azure Sponsorship 2	Deployments : 2 Succeeded		
Activity log	Subscription ID : b47e2e67-6ce6-430d-b41a-c4	39e1ff22fb		
R Access control (IAM)	Tags (change) : Click here to add tags			
🗳 Taos		*		
F Events	Filter by name Type == all O	ocation == all Add filter		
Cattions	Showing 1 to 7 of 7 records. Show hidden types	D	No grouping	\sim
Securgs	Name 1	Type 1.	Location 1+	
Quickstart		i per ce	Example 2	
Resource costs		Virtual network	East 052	
Deployments		Storage account	East US 2	
Policies	adtvm	Virtual machine	East US 2	
Properties	adfvm-ip	Public IP address	East US 2	
🗄 Locks	adfvm-nsg	Network security group	East US 2	
Export template	adivm239	Network interface	East US 2	
Monitoring	adfvm_OsDisk_1_80c54d02477c4b8290f64d835	ibeddc36 Disk	East US 2	
Insights (preview)				
Insights (preview) adfsalesdata				Å
Insights (preview) adfsalesdata swarge scourt Saarch (Ch1+0) search (Ch1+0)	≧ Open in Explorer → Move ○ Refresh	Delete G Feedback		Я
Insights (preview) adfsalesdata stange scouet P Search (Cbr(+/)) Coverview	⊇ Open in Explorer → Move ① Refresh Classic alerts in Azure Monitor is announced to retire i	Delete Oc Feedback 2021, it is recommended that you upgrade your classic alert rules to retain alerting fi	unctionality with the new alerting platform. For more information,	А
Insights (preview) adfsalesdata bioreget account Ø Search (Cbr(+/)) @ @ Overview Activity log	Gpen in Explorer → Move Refresh Classic alerts in Azure Monitor is announced to retire is see Continue alerting with ARM storage accounts: 0 ^o	Delete Geredback 2021, it is recommended that you upgrade your classic alert rules to retain alerting fu	unctionality with the new alerting platform. For more information,	Я
Insights (preview) disreger sclovel Search (Ctrl+/) Overview Conview Activity log A Access control (AM)	 ≧ Open in Explorer → Move ◯ Refresh Classic alerts in Azure Monitor is announced to retire is see Continue alerting with ARM storage accounts: of 	Delete Or Feedback 2021, it is recommended that you upgrade your classic alert rules to retain alerting fu	unctionality with the new alerting platform. For more information,	А
Insights (preview) adfsalesdata Servey account Servey account Actively log Access control (IAM) Tass Tass Tass Tass Tass Tass Tass Tass Tass Tass Tass	Open in Explorer → Move () Refresh Classic alerts in Azure Monitor is announced to retire is see Continue alerting with ARM storage accounts. () Resource group (change) : adf_lab	Delete Or Feedback 2021, it is recommended that you upgrade your classic alert nules to retain alerting fu Performance/Access tier : Performance/Access tier :	unctionality with the new alerting platform. For more information, Standard/Hot	Å
Insights (preview) Image assuest Image assumest	Copen in Explorer → Move C Refresh Classic alerts in Azure Monitor is announced to retire is ce Continue alerting with ARM storage accounts. D Resource group (change) : adf_lab Status : Primary: Available Location : East US 2	Delete Correction Performance/Access tier Replication	unctionality with the new alerting platform. For more information, Standard/Hot Locally-redundant storage (LRS) Sincand/2 (unexcal number v?)	¢
Insights (preview) adfsalesdata therape account Search (Ctrl+/) Access control (AM) Tags Access control (AM) Tags Data tanofer Data tanofer	Copen in Explorer → Move Refresh Classic alerts in Azure Monitor is announced to retire is ce Continue alerting with ABM storage accounts: of Resource group (change) : adf_lab Status : Primary: Available Location : East US 2 Subscription (change) : East US 2	Delete Co Feedback Total that you upgrade your classic alert rules to retain alerting fu Performance/Access tier : Replication : Account kind :	unctionality with the new alerting platform. For more information, Standard/Hot Locally-redundant storage (LRS) StorageV2 (general purpose v2)	¥
Insights (preview) Constraints (preview)	Copen in Explorer → Move Refresh Classic alerts in Azure Monitor is announced to retire is xee Continue alerting with ARM storage accounts: 0* Resource group (change) : adf_lab Status : Primary: Available Location : East US 2 Subscription (change) : Microsoft Azure Sponsor Subscription ID : b4726467-6c64-30d-b4	Delete Original State	unctionality with the new alerting platform. For more information, Standard/Hot Locally-redundant storage (LRS) StorageV2 (general purpose v2)	Å
 Insights (preview) adfsalesdata therape account Search (Ctrl+/) « Overview Activity log Access control (IAM) Tags Diagnose and solve problems Data transfer Events Storage Explorer (preview) 	Copen in Explorer → Move C Refresh Classic alerts in Azure Monitor is announced to retire is cecontinue alerting with ARM storage accounts: C ² Resource group (change) : adf_lab Status : Primary, Available Location : East US 2 Subscription (change) : Microsoft Azure Sponsor Subscription (change) : D472267-6c6-4304-b4 Tags (change) : Click here to add tags	Delete Original State	unctionality with the new alerting platform. For more information, Standard/Hot Locally-redundant storage (LRS) StorageV2 (general purpose v2)	я
Insights (preview) adfsalesdata Barge estudie Search (Cbr+/)	Copen in Explorer → Move C Refresh Classic alerts in Azure Monitor is announced to retire is see Continue alerting with ABM storage accounts. C ² Resource group (change) : adf_lab Status : Primary: Available Location : East US Subscription (change) : Microsoft Azure Sponsor Subscription (D : bd7e2e67-6ce6-430d-bd Tags (change) : Click here to add tags	Celete Celete Feedback 2021. It is recommended that you upgrade your classic alert rules to retain alerting fu Performance/Access tier : Replication Account kind in- stylp 2 Ia-c439e1ff22fb	unctionality with the new alerting platform. For more information, Standard/Hot Localy-redundant storage (LRS) StorageV2 (general purpose v2)	я
Insights (preview) Consequences of the second tempe account Conservation Second tempe account Second tempe account Access keys Tags Access keys	Copen in Explorer → Move	Delete Feedback r2021. It is recommended that you upgrade your classic alert rules to retain alerting fu Performance/Access tier : Replication : Account kind : Stylp 2 Tac-d39e1ff22fb R Tablac Tablac Tablac	unctionality with the new alerting platform. For more information, Standard/Hot Localy-redundant storage (LRS) StorageV2 (general purpose v2)	\$
Insights (preview) Constraints (preview)	Copen in Explorer → Move C Refresh Classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire a classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor is announced to retire alerts classic alerts in Acure Monitor alerts classic alerts alerts	Celete Feedback Performance/Access tier : Replication : Account kind : Severies SMB file shares File shares Celetary Celetary Celetary C	unctionality with the new alerting platform. For more information. Standard/Hot Localy-redundant storage (LRS) StorageV2 (general purpose v2) ge ge ge Cucues Effectively scale apps according	Å
Insights (preview) Constraints (preview)	[™] Open in Explorer → Move [™] Refresh [™] Classic alerts in Azure Monitor is announced to retire is ce Continue alerting with ARM storage accounts: 0° [™] Classic alerts in Azure Monitor is announced to retire is ce Continue alerting with ARM storage accounts: 0° Resource group (change) : adf_lab Status : Primays: Available Location : East US 2 Subscription (change) : Microsoft Azure Sponsor Subscription (change) : Click here to add tags <u>Containers Scalable-cost-effective storage for unstructured data </u>	Delete Original Second Secon	Inctionality with the new alerting platform. For more information, Standard/Hot Locally-redundant storage (LRS) .StorageV2 (general purpose v2) ge ge	چ پ ا

• Now we are going to go back to the virtual machine to connect it to the rest of the work from within the virtual machine. For doing that we are going to click the connect button.

≡ Microsoft Azure	, Sea	rch resources, services, a	nd docs (G+/)	-		D 🕼 🗘 🏟	? 😊
All services > CreateVm-Microso	ftWindows	Server.WindowsServer-20	01-20200509210531 Overview > adfvm > a	df_lab > adfvm			
adfvm Virtual machine							
,O Search (Ctrl+/)	~	🖉 Connect ▷ Sta	art 🦿 Restart 🔲 Stop 😥 Capture 🚺	Delete 🕐 Refresh			
Qverview		RDP (Im)	b		Azure Spot	: N/A	
Activity log	-	SSH	ng		Public IP address	: 137.116.94.15	
		Partion	JS 2		Private IP address	: 10.1.9.4	
R Access control (IAM)			- microsoft Azure Sponsorship 2		Public IP address (IPv6)	: •	
Tags		Subscription ID	: b47e2e67-6ce6-430d-b41a-c439e1ff22ft	b	Private IP address (IPv6):-	
Diagnose and solve problem	s	Computer name	: (not available)		Virtual network/subnet	: adf_lab-vnet/default	
Settings		Operating system	: Windows		DNS name	: Configure	
A New Mar	- 11	Size	: Standard D4 v3 (4 vcpus, 16 GiB memory	n			
Connect		Tags (change)	: Click here to add tags				



• Then we are going to select the RDP option and download the RDP file. Once it is downloaded we will open the file and we will log in the virtual machine with the user and password that we created previously.

Virtual machine		
	Λ To improve security, enable just-in-time access on this VM. $ ightarrow$	
Overview	RDP SSH BASTION	
Activity log		
Access control (IAM)	Connect with RDP	
Tags	the RDP file.	
Diagnose and solve problems	IP address *	
	Public IP address (137.116.94.15)	
Settings		
A Networking	Port number *	1
	3389	
S Connect		
S Disks	Download RDP File	

- Now that we are log in our virtual machine we are going to access to the azure portal to do the rest of the work from there. Now we are going to create and configure the azure data factory environment from creating our ETL process.
- For doing that we are going to select the "more services" option and then search for the data factories resource.

zure servic	es								
Create a resource	Virtual machines	Storage accounts	Resource groups	دی المی المی المی المی المی المی المی الم	Azure Active Directory	Help + sup	port Cost Management	Subscriptions	→ More services
All services	,								
Overview		🔛 Data fa	octories			SC Key	QL databases www.ds: Stretch Database		
Categories		📓 Azure I	Database for MySQ	L servers		No Da	ata Catalog		
All		🕑 Data La	ake Analytics			Da	ata Lake Storage Gen1		
General		💠 Data Be	ox			📑 Da	ata Shares		

- Now we are going to create and configure the new data factory resource in order to create our ETL process.
- We will need to configure:
 - **Name:** Name of the data factory
 - Version: Version of the data factory
 - **Subscription:** Type of subscription



- **Resource group:** The resource group which it belongs to
- **Location:** Where the SAP HANA instance is at, so that would be within the same data center

New data factory

Name *	
adfabforsap	×
Version ()	
V2	~
Subscription *	
Microsoft Azure Spansorship 2	~
Resource Group *	
adf_lab	3
Create new	
Location * ①	
(US) East US 2	~
Enable GIT 🛇	

 If we go to the resources page we can see that the data factory was created successfully. Now we are going to create the pipeline for our ETL process. For doing that we need to click the "author and monitor option" and then select the create pipeline button.

,Ϙ Search (Ctrl+/) «	Delete				
Overview	Resource group (change) :	adf_lab		Туре	: Data factory (V2
Activity log Access control (IAM) Tags Diagnose and solve problems ettings	Status : Location : Subscription (change) : Subscription ID :	Succeeded East US 2 Microsoft Azure Sponsorship 2 b47e2e67-6ce6-430d-b41a-c439	e1ff22fb 1	Getting start	ed : Quick start
Cocks	Documentation		Author & Monitor		



• Once we create the pipeline we need to configure the connection to SAP HANA. For doing that, we need to create a gateway to the SAP environment.



Factory Resources	* *
,P Filter resources by name	+
 Pipelines 	1
 GD pipeline1 	
Datasets	0
Data flows	0



• Once we click the connections button we are going to go to the integration runtimes tab to add another integration runtimes.



 Now we need to select which type on integration runtimes we want. In our case we are going to select Azure, Self-Hosted



• Next thing we must do is to give a name to our integration runtime.



Integration runtime setup

Private network support is realized by installing integration runtime to machines in the same on-premises network/VNET as the resource the integration runtime is connecting to. Follow below steps to register and install integration runtime on your self-hosted machines.

Name *	0
integrationRuntimetap	
Description	
Enter description here	
	la la
Туре	
Self-Hosted	

• The next window that appears is for making the installation of the integration runtime in our case we are going to make the installation with the option on express that will automatically make the installation. The reason that we are doing it within the virtual machine is so that it just stores automatically into it.

Integrati	on runti	me setup		
Settings	Nodes	Auto update	Sharing	
Install integ Authenticat	tion run	time on Window	s machine or add further nodes using the	
Name				0
integration	Runtimesa			
Option 2:	Manual	setup		
Step 1: Do	wnload and	d install integrati	on runtime	
Step 2: Use	this key to	register your in	tegration runtime	
NAME	AUT	HENTICATION KEY		
Key1	IR	@890481f4-dbdd-4	lba3-91b5-256ae17bece8@adflabforsap@eu2@FyJ3	D O
Key2	IR	@890481f4-dbdd-4	lba3-91b5-256ae17bece8@adflabforsap@eu2@Zo4\	D O

• Once the integration runtime is downloaded we click the install button to install it. It is important to know that an integrated runtime could only be associated to one azure data factory, this means that the integration runtime itself could be connected to many data sources but only to one Azure Data Factory.





• Now we'd copy the driver in



	> This	s PC > Desktop > SAP_HANA_CLIENT > :	SAP_HANA_CLIENT >			v ö s	earch SAP_HANA_CLIENT	
Ouldtan		Name ^	Date modified	Type	Size			
Quick access		client	5/10/2020 2:32 AM	File folder				
Desktop		excel	5/10/2020 2:33 AM	File folder				
Downloads	\$	instructime	5/10/2020 2:33 AM	File folder				
Documents	\$	licenses	5/10/2020 2:33 AM	File folder				
Pictures	*	D hdbclientreg	7/25/2019 12:23 PM	Application	43 KB			
This DC	_	D hdbinst	7/25/2019 12:23 PM	Application	43 KB			
This PC		hdbsetup	7/25/2019 12:23 PM	Application	43 KB			
Network		D hdbuninst	7/25/2019 12:23 PM	Application	43 KB			
		LABELASC	7/25/2019 12:23 PM	ASC File	1 KB			
		vcruntime140.dll	7/25/2019 12:23 PM	Application extension	84 KB			



AUTOMATATION ETL PROCESS

Once we have set up the necessary configuration we can begin to automate our process with Azure Data Factory. What we are going to do is copy some data from SAP to Azure Blog storage.

• The next step is to create a new dataset and then select the SAP HANA conector.



• Then we are going to create an SAP connection

SAP HAI HANA SapHan	IA aTable1			
General Connection	Parameters			
Linked service *	Select	\sim	HinNew	
Table	None	\sim	🕐 Refresh	60 Preview data



- For the configuration of the new connection, we need to file in the different fields:
 - **Name:** The name of the connection
 - **Connection via integration runtime**: The integration runtime that we have created previously
 - \circ $\,$ Server name: The name of the server that we want to connect to.
 - Authentication type: The type of authentication that you want to establish
 - Username
 - o Password

	·					
Name *						
Бармалат						
Description						
	la de la della d					
Connect via integration runtime *	0					
integrationRuntimesap	~					
Edit integration runtime						
Server name *						
52 2.251:31015						
Ratic authentication	~					
basic addrenication						
User name *						
DEMOUSER						
Password Ature Key Vault						
Password *						
Additional connection properties						
+ New						
Annatations						
Annotations						
- New						
Advanced ①						
	Connection successful					
Create	STest connection Cancel					

• The next step is to create the pipeline for our process.







• Once the pipeline is created, since we want to copy the data from SAP to Azure Blog Storage, we will use the activity called Copy data, which is in the Move and Transformation section.



• Now we are going to configure our component. For the General tab we are going to fill in the name of our component.

Copy data Sephanatoazureblog Image: Source 1 Sink 1 Mapping Settings User properties Name * Sephanatoazureblog Learn more Image: Settings User properties Name * Secretaria Description Timeout 7.00.0000 Retry 0 Retry interval		0		
P + - □ □ □ □ □ <td< th=""><th></th><th>Copy data</th><th></th><th></th></td<>		Copy data		
P + → <td< th=""><th></th><th>Sanhanatoazurehlog</th><th></th><th></th></td<>		Sanhanatoazurehlog		
P + - A ™ D N T + - A Mapping Settings User properties Seneral Source ¹ Sink ¹ Mapping Settings User properties Mame * Saphanatoszureblog Learn more I Description imeout 7.00.00.00 0 0 0 0 0		- sapiranatoazurebiog		
P + - Seneral Source ¹ Sink ¹ Mapping Settings User properties Iame * saphanatoazureblog Learn more I pescription imeout 7.00.00.00 imeout 7.00.00.00 ietry 0 ietry	L			
P + eneral Source ¹ Sink ¹ Mapping Settings User properties lame * saphanatoazureblog learn more Imeout 7.00.00000 O etry O o O				
P + - eneral Source ¹ Sink ¹ Mapping Settings ume * saphanatoazureblog escription				
0 + - eneral Source ¹ Sink ¹ Mapping Settings ame * saphanatoazureblog escription imeout 7.00.00.00 etry 0 0				
0 + - eneral Source Sink Mapping Settings user properties ame * saphanatoazureblog escription meout 7.00.00.00 o etry 0 ame value				
ame * saphanatoazureblog escription meout 7.00.00.00 etry 0				
P + - eneral Source ¹ Sink ¹ Mapping Settings ume* saphanatoazureblog escription imeout 7.00.00.00 etry 0 o etry 0				
P + - ieneral Source ¹ Source ¹ Sink ¹ Mapping Settings User properties Iame * saphanatoazureblog Learn more * imeout 7.00.00.00 etry 0 etry 0 and				
ieneral Source ¹ Sink ¹ Mapping Settings User properties lame * saphanatoszureblog Learn more I bescription				
lame * <u>saphanatoazureblog</u> Learn more ☑ lescription imeout 7.00.00.00 etry 0 etry 130 0	° + - 6			
lame * <u>saphanatoazureblog</u> Learn more ☑ Description imeout 7.00:00:00 tetry 0 tetry interval 30	⊳ + − €	Sink ¹ Manning Settings User	vopartias	
Pescription imeout 7.00.00.00 letry 0 letry 1 30 C	◦ + − € ieneral Source ¹	B 🚾 🛱 💽 👫 📲 Sink ³ Mapping Settings User p	properties	
etry 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	P + - E	Sink ¹ Mapping Settings User p	Droperties	
imeout 7.00:00:00 © etry 0 0 © etry interval 30 ©	o + - E	Sink ¹ Mapping Settings User p	Learn more	
imeout 7.00:00:00 © etry 0 © etry interval 30 ©	P + - E ieneral Source ¹ lame * rescription	Sink ¹ Mapping Settings User p	Learn more	
imeout 7.00:00:00 O etry 0 O etry interval 30 O	P + - E ieneral Source ¹ lame * rescription	Sink ¹ Mapping Settings User p	Learn more	
etry 0 0 4	P + - E	Sink ³ Mapping Settings User p	Learn more	
etry interval 30 O	P + - E ieneral Source ¹ tame * tescription imeout	Sink ³ Mapping Settings User p	Learn more	
30 O	P + - E seneral Source ¹ Hame * Description timeout tetry	Sink ³ Mapping Settings User p	Concepties	
	P + - E seneral Source ¹ Hame * Description simeout tetry	B [∞] [2] [3] II II II Sink ³ Mapping Settings User p saphanatoazureblog 7.00:00:00 0	Concepties	
	P + - E ieneral Source ¹ lame * Description imeout letry tetry interval	Sink ¹ Mapping Settings User p	Concepties	2
	P + - E General Source ¹ Name * Description Timeout Retry Retry interval Ensure output	Sink ¹ Mapping Settings User p	Concepties	2



• For the **source** tab we are going to select the SapHanaTable1 that we have created previously when we created our dataset, and we are going to click the open option to select a specific table of our dataset.

X Connections	I SapHanaTable1 I DipleIne2 I
Activities *	ĸ 🗧 Save as template 🗸 Validate ▷ Debug 💿 Add trigger
	Copy data
▲ Move & transform	Saphanatoazureblog
🗣 Copy data	∎ □ □ • ↔
Data flow	
Azure Data Explorer	
Azure Function	
Batch Service	
Databricks	
Data Lake Analytics	
Ø General	
P HDInsight	
 Machine Learning 	General Source Sink Mapping Settings User properties
, maaning scanning	
	Source dataset * SapHanaTable1 V Ø Open + New 66 Preview data
	Use query
	Partition option None Physical partitions of table Dynamic range D
	Packet size (KB)
	Additional columns O + New
X Connections	I SapHanaTable1
SAP SAP HANA	
HANA SapHanaTa	ble1
General Connection	Parameters
	Connection successful
Linked service *	SapHana1 V Ø Test connection Ø Open + New
Table	DEMODRIDAILY SALES
(and the	Edit



• For the **sink** tab we are going to create a new sink that it's where we are going to copy our data from SAP. First we need to create the dataset where we are going to copy our data. In this case its going to be de Azure blob storage.



• And them we need to select the format type of our data







- For the CSV format type, we must set its properties:
 - Name
 - **Linked Service**: We are going to create a new linked service that its going to be the Azure Blob Storage where we want to copy our
 - data.

Set properties	
Name	
sapdailysales	
Linked service *	
Select	~
Filter	
Select	
+ New	
in .	

- For the configuration of the new linked service, we need to fill in:
 - Name
 - Connect via integration runtime
 - Storage account name

hosted integration runtime is higher than version 4.0 if connecting via self- integration runtime.	hosted
Name *	
AzureBlobStorage1	
Description	
Connect via integration runtime *	C
AutoResolveIntegrationRuntime	~
Authentication method	
Account key	~
Connection string Azure Key Vault	
Account selection method From Azure subscription Enter manually	C
Azure subscription	0
Microsoft Azure Sponsorship 2 (b47e2e67-6ce6-430d-b41a-c439e1ff22fb)	\sim
Storage account name *	
adfsalesdata	\sim
Additional connection properties	
+ New	
Test connection	
To linked service To file path	
Annotations	
+ New	
📀 Connection succes	sful

• To finish configuring our component we need to set the las properties for the linked service that we have just created.



Set properties			
Name			
sapdailysales			
Linked service *			
AzureBlobStorage1			\sim
Edit connection			
File path			
salesdata	/ Directory	/ File	Browse 🗠
First row as header			
Import schema			
From connection/sto	ore OFrom sample	e file 🔵 None	
Advanced			

• The last step is to publish our process so we can test it properly.



• Once we have published it, we can execute it directly from Add trigger.



 And if we go to the monitor tab we can see how our process has been executed correctly and therefore the data has been copied from SAP to Azure blob storage.



Micro	osoft Azure Data Factory	adflabforsap					ي 😓	P ©	?
»	 Dashboards Dipeline runs A Trigger runs 	Pipeline runs Time : Last 24 hours (5/9/20)	2:17 AM - 5/10/20 2:17 AM)	Time zone : Monrovi	a, Reykjavik (UTC+0)	Runs : Latest runs	List	Gantt	
0	Integration runtimes Alerts & metrics	Showing 1 - 1 items PIPELINE NAME pipeline2	RUN START 14 5/10/20, 2:43:29 AM	DURATION 00:00:07	TRIGGERED BY Manual trigger	status In progress	PARAMETERS	A	NNOTATIO

CONECTOR ODP

The new SAP ODP connector leverages SAP Operational Data Provisioning (ODP) framework, which is an established best practice for data integration within SAP landscapes.

ODP provides access to a wide range of sources across all major SAP applications and comes with built-in CDC (change data capture) capabilities.

Most of the connectors can only extract data in batches, where each batch treats old and new data equally without identifying data changes . This extraction mode isn't optimal when dealing with large data sets, such as tables with millions or even billions of records, that change often. To keep your copied SAP data fresh, frequently extracting it in full is expensive and inefficient.

This connector can extract only data changes using CDC capabilities provided by SAP systems. It can connect to all SAP systems that support ODP, such as ECC, S/4HANA, BW, and BW/4HANA, directly at the application layer or indirectly using SAP Landscape Transformation (SLT) replication server as a proxy.

The connector can fully or incrementally extract SAP data that includes not only physical tables, but also logical objects created on top of those tables, such as Extractors or ABAP Core Data Services (CDS) views, without watermarking.

SAP ODP connector can extract various data source types, such as:

- SAP extractors, originally built to extract data from SAP ECC and load it into SAP BW
- ABAP CDS views, the new data extraction standard for SAP S/4HANA
- InfoProviders and Info Objects in SAP BW or BW/4HANA
- SAP application tables, when using SLT replication server as a proxy



CONNECT AND COPY SAP HANA DATA WITH AZURE DATA FACTORY

Prerequisites:

- Configure SAP systems to use <u>SAP's Operational Data Provisioning (ODP)</u> <u>framework.</u> (Acts as a central data persistency layer and supports extraction and replication scenarios for various target SAP applications)
- Become familiar with Data Factory concepts such as integration execution environments, linked services, datasets, activities, data flows, pipelines, and triggers.
- Configure a self-hosted integration execution environment to be used for the connector.
- Configure an SAP CDC linked service.
- Debug SAP CDC connector issues by sending logs from the self-hosted integration runtime environment to Microsoft.
- Become familiar with monitoring data extractions in SAP systems.
- Configuring SAP systems to use the SAP ODP framework

Requirements for SAP systems:

- To support ODP, run SAP systems on SAP NetWeaver 7.0 SPS 24 or later. (Transferring Data from SAP Source Systems via ODP (Extractors)).
- To support full SAP Advanced Business Application Programming (ABAP) Core Data Services (CDS) extractions via ODP, run SAP systems on NetWeaver 7.4 SPS 08 or later. To support SAP ABAP CDS differential extractions, run SAP systems on NetWeaver 7.5 SPS 05 or later. <u>(Transferring Data from SAP Systems via ODP</u> (ABAP CDS Views)).
- 1521883: to use ODP API 1.0
- 1931427: to use ODP API 2.0 supporting SAP hierarchies
- 2481315: to use ODP for the extraction of data from SAP source systems to BW or BW/4HANA systems

SAP user configuration

- Data extractions via ODP require a correctly configured user in SAP systems. The user must be authorized to invoke the ODP API via Remote Function Call (RFC) modules. For more information, please refer to these SAP Support Notes:
 - o 2855052: to authorize the use of the ODP API.
 - 460089: to authorize the invocation of ODP RFCs



Configuration of SAP Data Origins

To extract data sources, make sure the following requirements are met:

- Ensure that the data sources are activated in the SAP source systems. This
 requirement only applies to data sources provided by SAP or its partners. Data
 origins created by customers are activated automatically. If data sources have
 been or are being extracted by SAP BW or BW/4HANA, the data sources are
 already activated. (Installing BW Content Data Sources)
- Ensure that the data sources are released for extractions via ODP. This requirement only applies to data sources that are created by customers. Data sources delivered by SAP, or its partners are automatically released. For more information, see these SAP Support Notes:
 - 1560241: to release data sources for the ODP API.

SAP Landscape Transformation Replication Server Configuration

- SAP Landscape Transformation Replication Server (SLT) is a database triggerenabled CDC solution that can replicate SAP application tables and simple views in near real-time. You can use SLT as a proxy for ODP data extraction. To use SLT as a proxy, complete the following steps:
 - Install NetWeaver 7.4 SPS 04 or later and the DMIS 2011 SP 05 snap-in on the replication server. For more information, see Transferring data from SLT using operational data provisioning.
 - Run the SAP Landscape Transformation Replication Server Cockpit (LTRC) transaction code on the replication server to configure SLT:
 - Under Specify Source System, type the RFC destination that represents the SAP source system.
 - Under Specify Target System, complete these steps:
 - Select RFC Connection.
 - Under Scenario for RFC communication, select Operational Data Provisioning (ODP).
 - Under Queue Alias, type the queue alias to be used to select the context for data extractions via ODP in Data Factory. Use the format SLT-<your queue alias>.

Specify General Data	Specify the relevant information in order to connect to the target system				
Specify Source System	Target System Connection Details				
Specify Target System Specify Transfer Settings Settings	RFC Connection	ODB Connection			
	Scenario for RFC Communication	Operational Data Provisioning (ODP)			
Wevew and create	RFC Destination	NONE			
	Queue Alas	ECCNSP			



Setting up a self-hosted integration runtime environment for the SAP CDC Connector

Creating and managing a self-hosted integration runtime environment

• In Azure Data Factory Studio, create and configure a self-hosted integration runtime environment. For downloading the self-hosted integration runtime environment, you can go to this page

Download and install the SAP .NET Connector

- Download the latest version of the <u>64-bit SAP .NET Connector (SAP NCo 3.0)</u> and install it on the machine running the self-hosted integration runtime environment.
- In the Optional Installation Steps dialog box, select Install Assemblies in GAC, and then select Next.

🛃 SAP .Net Connector 3.0 for .NET 4	.0 on x64		-		\times
Optional setup steps					
Installing assemblies to GAC allows differ assemblies. Registering WMI provider al Connector 3 so that the WMI objects ca	ent applications on lows applications to n be queried or mor	the target comp publish WMI ob itored using WN	utertos jectspr 11 infras	share the ovided by tructure.	.NET
 None Install assemblies to GAC 					
O Register WMI provider and install	assemblies to GAC				
					_
	Cancel	< Back		Next	>

Adding a network security rule

If the SAP system is on an Azure virtual machine, to add the rule you must:

- Source IP addresses/CIDR ranges: Set Source IP Addresses or CIDR ranges to the IP address of the self-hosted integration runtime environment machine.
- Set Destination Port Ranges in 3200,3300.

Checking Connectivity

On the machine running the self-hosted integration runtime environment, run the following PowerShell cmdlet to ensure that you can connect to the SAP systems:

• Test-NetConnection <SAP system IP address> -port 3300



Editing hosts files

Edit the hosts file on the machine that is running the self-hosted integration runtime environment to add the SAP IP addresses to the server names

 On the machine running the self-hosted integration runtime environment, edit C:\Windows\System32\drivers\etc\hosts to add mappings of the SAP system IP addresses to the server names. For example: # SAP ECC xxx.xxx.xxx.sapecc01 # SAP BW yyy.yyy.yyy.yyy sapbw01 # SAP SLT zzz.zzz.zzz sapnw01

Configuration of a linked service and a data set for the SAP CDC Connector

Configuring linked Services

 In Azure Data Factory Studio, go to the Data Factory Administration center. From the Connections menu, select Linked Services. Select New to create a linked service.



• Under New Linked Service, search for SAP. Select SAP CDC and then choose Continue.



D sap				
All Azure Databa	ase File	Generic protoc	ol NoSQL	Services and apps
SAP		SAP BW		SAP
SAP BW Open Hub		SAP BW via MDX	SA	P CDC
C4C		SAP		SAP
		ECC		HANA
SAP Cloud For Custon	ner	SAP ECC		SAP HANA
CAD				

- Set the properties of the linked service:
 - **Name**: Type a unique name for the bound service.
 - **Connect via integration runtime**: select the configured integration runtime environment.
 - Server name: name of the assigned server for the SAP system.
 - **Subscriber name**: Unique name to register and identify this Data Factory connection as a subscriber that consumes data packets that are generated in the Operational Delta Queue (ODQ) by the SAP system.

SapCDCLinkedService		
Description		
	h	
Connect via integration runtime * ①		
Select	~	
Logon type		
Application server Group		
Server name *		1
System number *		L
Client ID *		L
Language		
SNC mode		1
On Off		
User name *		
Password Azure Key Vault		
Password Azure Key Vault Password *		
Password *		
Password *		
Password * Password * Subscriber name MY_ADF		

Configuring source data

- In Azure Data Factory Studio, go to the Data Factory Authoring centre.
- In Factory Resources select New Data Set.





- Under New Data Set, search for SAP.
- Select SAP CDC and then choose Continue.



- Set Properties: enter a name for the SAP CDC linked service data source and under Linked Service, select the drop-down list, then select New.
- Select the SAP CDC linked service for the new source data set and set the rest of the properties for the linked service:
 - \circ $\:$ Under Connect via integration runtime, select the configured integration runtime environment.



- Under ODP context, select the ODP data extraction context. Here are some examples:
 - To extract ABAP CDS views from S/4HANA, select ABAP_CDS.
 - To extract InfoProviders or InfoObjects from SAP BW or BW/4HANA, select BW.
 - To extract SAP extractors from SAP ECC, choose SAPI.
 - To extract SAP application tables from SAP source systems using the SLT replication server as a proxy, select SLT_<queue alias>.

If you want to extract tables from SAP applications, but do not want to use SAP Landscape Transformation Replication Server (SLT) as a proxy, you can create SAP extractors using the RSO2 transaction code or Core Data Services (CDS) views with the tables. Then extract the tables directly from the SAP source systems via a SAPI or an ABAP CDS context.

• Under ODP Name, in the selected data extraction context, select the name of the data source object to be extracted. If you connect to the SAP source system via SLT as a proxy, the Data Preview feature is currently not supported.

Name			
SAP_ECC_FI_A	counts_Receivable		
Linked service			
SAP_CDC_Link	ad_Service	~	0
Connect via int	gration runtime * ①		
integration	untime1	\checkmark	Ø
Context			
SAPI		~	Ü
Edit			
Object name			
None		~	U
0FI_AR			
OFI_AR_4			
0FI_AR_5			
OFI_AR_6			

Data transformation with the SAP CDC connector

Mapping Data Flow Instance Mapping Properties

- To create a mapping data flow using the SAP CDC connector as the source, complete the following steps:
 - o In ADF Studio, go to the Pipeline section of the Creation centre





- \circ $\:$ Select the ... button to open the Pipeline Actions drop-down menu.
- Select the New Pipeline item.
- Activate the debug mode using the Debug Data Flow button in the top bar of the data flow canvas.



• In the mapping data flow editor, select Add source.

Validate Data flow debug				🖗 dataflow1	» å∆ d
		Data flow debug	Data	🗸 Validate 🌘	~
Add Source	1	Add Source	Source	Add	
Add Flowlet	1	🛼 Add Flowlet			l

• In the Source Configuration tab, select a prepared SAP CDC dataset or select the new button to create a new one. Alternatively, you can also select Inline in the Source Type property and continue without defining an explicit dataset.





🔛 Data Factory 🗸 🛁 Valid	date all 📋 Publish all 2
» 🖧 dataflow1 ●	
Validate Data flow	debug
Columna: 0 total	+
Add Source	
Source settings Source opt	tions Projection Optimize Inspect Data preview
Source settings Source op Output stream name *	source1
Source settings Source op Output stream name * Description	Itions Projection Optimize Inspect Data preview source1 Learn more [3] Import data from SapCdcResource C Reset
Source settings Source op Output stream name * Description Source type *	Import data from SapCdcResource Carton Control Contro Control Control Control Control Control Contro Control Con
Source settings Source op Output stream name * Description Source type * Dataset *	tions Projection Optimize Inspect Data preview source1 Learn more (3) Import data from SapCdcResource (2) Reset
Source settings Source op Output stream name * Description Source type * Dataset * Options	tions Projection Optimize Inspect Data preview source1 Learn more (?) Import data from SapCdcResource (?) Reset Import data from SapCdcResource (?) Reset Import data from SapCdcResource (?) Reset Import data from SapCdcResource (?) Open + Ne Import data from SapCdcResource (?) Open + Ne Import data from SapCdcResource (?) Open + Ne
Source settings Source op Output stream name * Description Source type * Dataset * Options	tions Projection Optimize Inspect Data preview source1 Learn more ? Import data from SapCdcResource
Source settings Source op Output stream name * Description Source type * Dataset * Options	tions Projection Optimize Inspect Data preview source1 Learn more ? Import data from SapCdcResource

- On the Source Options tab select:
 - Full on every run option if you want to load full snapshots on every run of the mapping data (pipeline) stream
 - Full on first run, incremental, if you want to subscribe to a source of changes from the SAP source system. In case of incremental loads, it is necessary to specify the keys of the ODP source object in the Key Columns property.

*	🊂 Data Factory 🗸 🖌 Validate all 📋 Publish all 🕕
↑	♦ 🖗 dataflow1 ● 🎟 SapCdcResource
	🗸 Validate 🚺 Data flow debug 🤣 🕞 Debug Settings
© ©	Columna: O total
۲	Source settings Source options Projection Optimize Inspect Data preview •
	Run mode * Full on every run
	Incremental changes only

• On the optimize tab we can see the optional values for configuring partitioning schemes.



Aggregate settings	Optimize Inspect	Data preview			
Partition option *	O Use current pa	rtitioning OS	ingle partition 🤇	Set Partitioning	
Partition type *	1 2 3 1 1 2 3		A } (1) } (2) Z } } (3)	Conditions 2 3	(ceys)
	Round Robin	Hash	Dynamic Range	Fixed Range	Key
Number of partitions *	20				

• On the Inspect tab we can see a view of the metadata of the data stream being transformed. You can see the number of columns, columns that have changed, columns that have been added, data types, column order, and column references. Inspect is a read-only view of the metadata.

Derived column	n's settings Optimize	Inspect Data Preview			🖵 Description
Output schema	Input schema				
Number of col	lumns New* 1	1	Updated [®] 2	Unchanged 4	Total 7
Order \$	Column ‡		Туре Ф	Updated \$	Based on \$
1	movie		also string		
2	title		als: string	•	title
3	genres		etc: string		
4	year		121 long	•	year
5	Rating		atic string		
6	Rotton Tomato		He string		
7	Rotten Tomato		121. long		Rotton Tomato



AUTOMATATION ETL PROCESS

• We will see an example using the conector to copy activity from a table in SAP followed by a template mapping data flow that processes the changes and updates it in an SQL table in Azure Synapse.

	dinter and the		₽ Search	
Vould you like to try preview updates	to Azure Data Fa	ctory Studio? Open settings to lea	rn more and opt in	
🕥 / 👯 main branch 🗸 🗸	🛁 Validate all	🔚 Save all 🕕 Publish		
Factory Resources	× « 0	D SapDataReplicatio × 🖧	sapcdcapplychange	
Y Filter resources by name	+	Activities 🛛 🗧 «	🔚 Saved 🗧 Save as template 🗸 Validate 🕨 Debug 🗸	Add trigger OD Data flow debug
Pipelines	24	₽ Search activities		
🖌 🗀 Data flow demo	1	> Move & transform		
000 SapDataReplicationSyn	apse	> Azure Data Explorer	N	If Condition
Demo copy activity	1	> Azure Function	L3*	
Test Pipelines	22	> Batch Service	Copy data	TriggerDataflow
Datasets	41	> Databricks	FetchSap	DeltaChanges True
 Data flows 	8	> Data Lake Analytics		False
CreateDimension		> General		No activities
CreateDimension_V3		> HDInsight		
CreateFacts V3		> Iteration & conditionals		
the createracts"Ap				
SalesOrdersDF		> Machine Learning		
SalesOrdersDF		> Machine Learning > Power Query		
ේ Createrace_vs තී SalesOrdersDF තී sapcdcapplychange තී sapcdcapplychange1		> Machine Learning > Power Query	Parameters Variables Settings Output	_
ේ Createristics 19 මී SalesOrdersDF මී sapcdcapplychange මී sapcdcapplychange1 මී TankerkoenigJoinPricesWitl	hStations	> Machine Learning > Power Query	Parameters Variables Settings Output	_
ේ Contract_vs හි SalesOrdersDF හි sapcdcapplychange හි sapcdcapplychange1 හි TankerkoenigJoinPricesWith හි TankerkoenigTest	hStations	> Machine Learning > Power Query	Parameters Variables Settings Output	-
 Createratescy vs Ĝ² SalesOrdersDF Ĝ³ sapcdcapplychange Ĝ³ sapcdcapplychange1 Ĝ³ TankerkoenigJoinPricesWitt Ĝ³ TankerkoenigTest Power Query 	hStations 0	> Machine Learning > Power Query	Parameters Variables Settings Output H New Delete Type	Default value
	hStations 0 2	> Machine Learning > Power Query	Parameters Variables Settings Output HNew Delete Name Type Context String ✓	Default value SLT-ECCNSP
^c Createriacs2,v3 ^c SalesOrdersDF ^c SalesOrdersDF ^c Sapcdcapplychange ^c TankerkoenigJoinPricesWitt ^c TankerkoenigJoinPricesWitt ^c TankerkoenigJoinPricesWitt ^c TankerkoenigJoinPricesWitt ^c TankerkoenigJoinPricesWitt ^c TankerkoenigJoinPricesWitt ^c TankerkoenigTest ^c TankerkoenigTest	hStations 0 2	> Machine Learning > Power Query	Parameters Variables Settings Output + New IDelete Image: String Image: String	Default value SLT-ECCNSP
^c Createriacs_vs ^c SalesOrdersDF ^c TankerkoenigJoinPricesWitt ^c TankerkoenigTest ^b Power Query ^b Templates	hStations 0 2	> Machine Learning > Power Query	Parameters Variables Settings Output + New IDelete Image: String Image: String	Default value SLT-ECCNSP EKKO
 Createrates_vs Ĝ² SalesOrdersDF Ĝ³ sapcdcapplychange Ĝ³ sapcdcapplychange1 Ĝ³ TankerkoenigJoinPricesWitl Ĝ⁴ TankerkoenigTest Power Query P Templates 	hStations 0 2	> Machine Learning > Power Query	Parameters Variables Settings Output + New Delete Type	Default value SLT-ECCNSP EKKO ['MANDT',"EBELN'] COPY FKO
 Createrates, vs Ĝ² SalesOrdersDF Ĝ³ sapcdcapplychange Ĝ³ sapcdcapplychange1 Ĝ³ TankerkoenigJoinPricesWiti Ĝ⁴ TankerkoenigTest Power Query P Templates 	hStations 0 2	> Machine Learning > Power Query	Parameters Variables Settings Output + Name Type	Default value SLT-ECCNSP ii EKKO ii ['MANDT',"EBELN'] ii COPY_EKKO ii SABPoren iii

🕥 / 👯 main branch 🗸 🖌 🗸	Jate all 层 Save all 🕕 Publish	0
Factory Resources *	« 0D SapDataReplication	1
√ Filter resources by name	+ 🔚 Saved 🗸 Validate 🔘 Data flow debug	
A Pipelines	24 🛛 sapcdcdeltachange getDataFlowKeyCoL. PartitionByKeyColu. MergeOperatik AlterRowForSynapse RemoveSystemColu. MergeOperatik	
Data flow demo DD SapDataReplicationSynapse	1 Constrained by the second se	
Demo copy activity	1 particular and the second	
Test Pipelines	22 Add Source	
Datasets	41	
 Data flows 	8	
ể∆ CreateDimension		
ể∆ CreateDimension_V3		
ể∆ CreateFacts_V3		
A SalesOrdersDF		
sapcdcapplychange1	Parameters Settings	
TankerkoenigJoinPricesWithStatic		
⊕⊃ TankerkoenigTest	+ New Delete	
Power Query	0 Name Type Default value	
Templates	2 folderPath and string V Enter expression Any 🗈	
	ter chimne the chine it is the average in the interview in the chine is the chine in the chine is the chine i	



- www.stratebi.com
- In the monitoring tool you can see that an initial load has been done

鎆 Dashboards Runs	Pipeline runs Triggered Debug Serun	S Cancel ✓ C Refrest	n 📰 Edit columns 🛛 🚺 🖬	Gantt						
00 Pipeline runs	Filter by run ID or name	Local time : Custom (6/22/22	, 8:00 PM - 6/22/22, 11:16 PM)	Pipeline nam	ne : All Status : All	Runs : Late	st runs	D	1 Empertur CE	
🖧 Trigger runs	Triggered by : All	filter ×						Copy filter	Export to CS	1.
Runtimes & sessions	Showing 1 - 1 items								Last refreshed 0 minu	utes a
ூ Integration runtimes	Pipeline name	Run start	Run end	Duration	Triggered by	Status	Error	Run	Parameters	
Data flow debug	SapDataReplicationSynapse	Jun 22, 2022, 8:02:52 pm	Jun 22, 2022, 8:04:45 pm	00:01:52	Manual trigger	Succeede	d	Original	1001	
Notifications										

• If we go to Azure synapse we can see that 16221 columns have been loaded into the table



- On the sap side, we can verify that this first pipeline run has initiated the Delta process. On this feed we will now be able to read change as we do in SAP purchasing
- Let's move on to the SAP purchasing aplication to create a new purchasing order. We will use an existing order made as a template for better simplicity.





	l 😒 💫 l	₽ 60 60	13	1000	🗖 🗖 I 🕜 🖳						
Create Durchase Order											
Create rurchase order											
Document Overview Off	Park é	0 1 📅 (B 1 b	rint Pr	eview Mes	ages 🔝 🔮 Personal Settin	ng Save As T	emplate Load	rom Templa	ite		
Description of the second seco	1007	IDES - GTS	PO	•	Vendor	1095 Lopez S.A.		Doc. date	22.06.2022		
Document Overview		Header		2							
CADEMEDORECT		E C Itm	AT	Material	Short Taxt	PO Quanthy	O C Daly	Data No	Price Our	Der	0
SAPONEFODEST		A 10	0 1	DTIL 7000	Elektropische Chine Fis Dr	PO Quanticy	200 0 16 0	1 2010	100 00 001	rei	0
				010-7000	Elektronische Chips für Br.	÷	SPC 0 10.0	1.2010	100.00 HAR	1	PC
									MAN		
urch.Doc.				_					MXN		
4500019689									MXN		
6 4500019689	÷								MXN		
									MICIN		
4500019689									MXN		
4500019689									MXN		
4500019689									MON		
4500019689			+++						MYN		
• 4500019689				-					Lon		
4500019689									PLAN		
4500019689	12								MXN		
4500019689	11 2										• •
4500019689							- from Malana		and and a second s		
- 4500019689 - 4500019689							erault values	Addi Pi	inning		
4500019689	-		-								
1000010002		Item	1	10] DTU-700	0, Elektronische Chips für Br	em 🔻 🔺 🔻					
4500019689		Materia	Data	Quantities	Weights Delivery Schedu	le Delivery	Invoice Co	nditions	Call-Off Texts	D	4 Þ
♦ 4500019689 ♦ 4500019689											
4500019689 4500019689 4500019689 4500019689						*			Address det	ale	
 4500019689 4500019689 4500019689 4500019689 4500019689 		Title						6	Nucleas deci	10	
 4500019689 4500019689 4500019689 4500019689 4500019689 4500019689 		Title		Nor	Vade						
 4500019689 4500019689 4500019689 4500019689 4500019689 4500019689 4500019689 		Title Name		Nev	/ York						
 4500019689 4500019689 4500019689 4500019669 4500019669 4500019669 4500019669 4500019669 		Title Name		Nev	/ York / York	Versen					
 ♣ 4500019689 		Title Name Street/Ho	use nu	Nev Nev Imber Bro	/ York / York idway	691					
 ¹ ¹		Title Name Street/Ho District	use nu	Nev Nev	/ York / York idway	691		A	ddress		
4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500119689 4 \$500019689 4 \$500019689 4 \$500019689 4 \$500019689 4 \$500019689		Title Name Street/Ho District Postal Coo	use nu je/Citv	Nev Nev Imber Bro	r York r York idway 12-1121 New York	691		A	ddress		

• Once we saved our new purchasing order we are going to trigger the process again to see what's happening.

🔿 / 👯 main branch 🗸 🖌 Valida	te all 🔚 Save all 📋 Publish		
Factory Resources × «	00 SapDataReplicatio ×	A sapcdcapplychange	
▼ Filter resources by name	Activities × «	🔚 Saved 😤 Save as template 🗸 Validate 🕨 Debug 🖌 🧏 Add trigg	er 🔘 Data flow debug
 Pipelines 2 	4 Search activities	Trigger not	Trigger on-demand run of the last published pipeline
🔺 🖹 Data flow demo	1 > Move & transform	New/Edit	🚽 🦆 TriggerDataflow
000 SapDataReplicationSynapse	> Azure Data Explorer	Copy data	True
Demo copy activity	1 > Azure Function	FetchSapDeltaChange	s 1 activities
Test Pipelines 2	2 > Batch Service		False D
Datasets 4	11 > Databricks		- no ocurrace
 Data flows 	8 > Data Lake Analytics		
CreateDimension	> General		
ể∆ CreateDimension_V3	> HDInsight		
ể∆ CreateFacts_V3	> Iteration & conditionals		
A SalesOrdersDF	> Machine Learning		
apcdcapplychange	> Power Query		

• We are going to see that for the copy activity the process has read only one row from the SAP system.





Details 🕐 Refres	h			
earn more on copy perform	mance details from here.			
Activity run id: 55e9b32a-f	552-4782-bf2f-1a01368d17cd			
SAP ODP	(Preview)	ceeded	Azure Data Region: W	a Lake Storage Gen2 est Europe
Data read: 🕕 🔓	652 bytes		Data written: ①	1,791 KB
Rows read:	1		Files written: ①	1
Peak connections: ①	2		Rows written: ①	1
			Peak connections: ①	1
Copy duration	00:00:10			
'hroughput: ①	65,536 byt	es/s		
\sim SAP ODP (Preview) \rightarrow	Azure Data Lake Storage Gen2			
Start time	Jun 22, 2022, 8	:21:54 pm		
Used parallel copies ①) 1			
\sim Duration	00:00:10			
Details		Working duration	Total duration	
O Queue	-	-	00:00:06	
Transfer ①	Listing source ① Reading from source ① Writing to sink ①	00:00:00 00:00:00 00:00:00	00:00:01	
Data consistency verifica	ation () Not verifie	d		
ata consistency vernica	Not verne			

• Finally, if we move to the synapse side to verify the data that we have retrieved we can see that now if we execute the query the result involves one more column, the purchase order that we have created previously.

Run	> Undo ~	(<u>1</u> F	Publish	പ് (Query plan	Connect
1	SELECT* FROM	[SAPDe	emo].[s	SAP_E	кко]	
Results	Messages					
			_			
View	Table	Chart		\mapsto	Export results	\sim
0.	rch					
₽ Sea	1.757 A.					



POSITIVE AND NEGATIVE ASPECTS REGARDING THE SAP CONECTION

Azure Data Factory is an interesting tool when connecting to SAP for automating a process. On the one hand, the configuration that must be followed to be able to use this type of connectors is somewhat tedious, but on the other hand, the two connectors presented above connect and interact directly with the SAP database.

In terms of data loading efficiency, it is worth mentioning that the ODP is much more efficient when working with large amounts of data, since for each trigger call it will only use the new data entered in the SAP tables to which it is connected instead of the entire table directly.

On the other hand, although it is more efficient, it must be taken into account that if the process to be created does not require much data load that could slow down the execution or if you want to make a simple data copy process, it would be better to use the usual SAP HANA connector since that de ODP, was released in July 2022, so it does not have as much documentation as the SAP HANA connector.

In general, in comparison to the rest of the tools, using Azure Data Factory to connect to SAP is more efficient in terms of data partitioning because it can load terabytes of data, very fast

Finally, keep in mind that to access to certain documentation related to SAP it is necessary to have an S-user account so sometimes it is difficult to access certain documentation.



6. **CONCLUSIONS**

Once we have seen the automation tools, we can see how there are big differences when choosing one tool or another to automate our process. In general terms, the choice of one tool or another will be dictated by the type of process we want to develop, as well as the type of licenses we have in our organization.

In general terms we can see how Azure Data Factory and Power Automate could be the most appropriate when automating processes. Since both tools have enough documentation to investigate and consult in case of any problem. However, Azure Data Factory would be more desirable in case the process to be automated had some complexity or followed a type of workflow that could not be created with Power Automate, since this tool allows to create much simpler flows.

On the other hand, Pentaho, although it is not a bad option, it is worth mentioning that it does not have extensive documentation and that it does not have a specific connector to connect to SAP HANA so it would be necessary to install the IT NOVUM plug-in, which is a tedious and complex process. In resume, it would be more advisable and, considered that Pentaho is not being used from the beginning, to use another tool.

Talend could be another option, and even if the specifications of the company require it when creating the process, it would not be necessary to buy the paid version of the tool. On the other hand, if the process to be automated is a simple process, it is not worth it because the installation of the tool is somewhat complicated and although there is a lot of documentation about it Talend is a complicated tool to use if you are new to it.

As for Power Automate, it should be noted that this tool is very intuitive and easy to use, but with it you cannot perform flows as complex as those you can perform for example in Talend. In this case to connect to SAP and extract information from the tables it would require a license because the connector is premium.

Finally, regarding Azure Data Factory, as we have seen there are several types of connectors with which we can connect to SAP, in this case we have explained the SAP HANA connector and the ODP connector for SAP. Using one or another connector would depend in part on the amount of data that would be handled in the tables that are in SAP, as well as the type of process that you want to carry out. Choosing this tool to automate processes or another tool will depend not only on the amount of data to be handled in the SAP tables, as ADF is capable of loading terabytes of data very fast but also in the frequency with which the process will be used because we must consider that with Azure Data Factory we pay for what we use.