

SAP Cloud Integration - An Overview, Best Practices, and Implementation Steps - Part 3

Deepak Kumar¹, Maha Bhageshwara Raju Kesaboina²

¹Wilmington, USA

Email: [deepak3830\[at\]gmail.com](mailto:deepak3830[at]gmail.com)

²Hyderabad, India

Email: [bhageshwar.raju\[at\]gmail.com](mailto:bhageshwar.raju[at]gmail.com)

Abstract: SAP CPI, also known as SAP Cloud Platform Integration is a cloud-based integration platform provided by SAP. It allows for connectivity, between applications, systems, and services both within and outside of an organization. With SAP CPI businesses can simplify the integration process. Optimize data flow and communication between software solutions. It facilitates the exchange of data across environments promoting interoperability and enhancing collaboration efficiency. Additionally, SAP CPI offers built-in integration iFlow and adapters to reduce the complexity of integration projects and speed up deployment timelines. Overall SAP CPI plays a role, in the SAP ecosystem by enabling the creation of agile business processes. In this paper, we will discuss SAP CPI's capabilities in connecting diverse applications, facilitating data flow, supporting different integration scenarios, and leveraging pre-built content and adapters. It will provide a step-by-step guide or overview of the fundamental steps involved in implementing SAP CPI. This paper will cover the initial setup, configuration, and the process of establishing connections between different systems. The aim is to offer a practical understanding for readers who might be considering or undergoing the implementation process. This paper will also address the monitoring aspect of SAP CPI, emphasizing the importance of keeping track of integrations and workflows. Insights into monitoring tools, key performance indicators, and best practices for ensuring the ongoing efficiency and reliability of integrated systems will be elaborated.

Keywords: SAP, SAP CPI, SAP Cloud Integration, SAP On-premise to SAP cloud integration

1. Introduction

SAP CPI – SAP Cloud Platform Integration is a powerful cloud-based platform that enables seamless integration between cloud applications and other SAP or non-SAP cloud or on-premises applications. SAP CPI uses Apache Camel in its integration framework. Apache Camel is an open-source framework that offers a variety of pre-built components and patterns for creating enterprise integration solutions. SAP CPI utilizes Apache Camel to make it easier to develop integration flows and to improve the platform's flexibility and extensibility.

ETL stands for Extract, Transform, Load, which refers to a process used in data warehousing and data integration.

API stands for Application Programming Interface. It's a set of protocols, tools, and definitions that enable different software applications to communicate with each other. APIs determine how software components should interact, allowing developers to access specific features or data from a service, library, or application without having to comprehend its internal workings.

SAP DS - SAP DS typically refers to "SAP Data Services," which is an enterprise-grade data integration and transformation software offered by SAP. SAP Data Services enables organizations to extract, transform, and load (ETL) data from various sources into target systems for analytics, reporting, and other business processes.

Pre-requisite: To understand this paper thoroughly prerequisite is SAP Cloud Integration – An Overview, Best Practices, and Implementation Steps – PART 1 and PART 2.

1) SAP Cloud Integration for Data Services

SAP CPI Standards, Best Practices - The purpose of this paper is to establish guidelines for integration developers who will be using the SAP Cloud Platform Integration Service to develop integrations for projects. In this paper, we will discuss the SAP Cloud Integration for Data Services business use case. Formerly known as HCI-DS, SAP Cloud Platform Integration for Data Services (CPI-DS) stands as a robust cloud-based integration tool by SAP. Its core purpose lies in facilitating batch and scheduled data integration between on-premise applications and the SAP cloud platform. CPI-DS bridges the gap between on-premise systems and the cloud, ensuring seamless data movement and synchronization across heterogeneous landscapes. This pivotal role enables smooth and efficient data flow, vital for supporting diverse business processes and analytics initiatives within organizations.

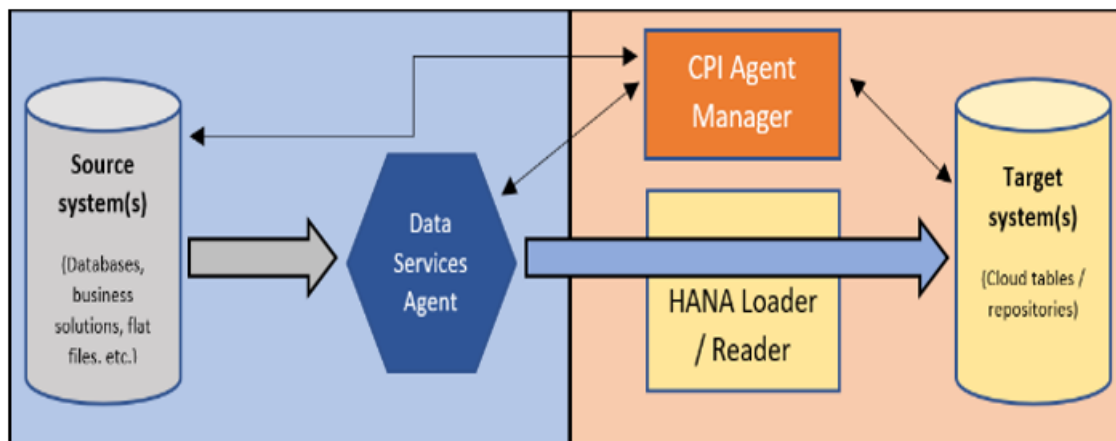
2) Positioning and context

SAP's Cloud Platform Integration (CPI), formerly known as HANA Cloud Integration or HCI, stands as a prominent cloud-based integration solution offered by the Walldorf-based giant. Unlike the on-premise alternative, Process Orchestration (PO), CPI encompasses multiple scenario-specific applications aimed at synchronizing data across an enterprise's system landscape. When customers invest in modern SAP business applications, they often receive bundled CPI applications, facilitating integration into the existing IT architecture. For instance, SAP's Integrated Business Planning (IBP) includes CPI-DS. However, CPI-DS isn't exclusive to IBP; it plays a significant role in scenarios involving S/4HANA, SuccessFactors, C4C, or BW, among others. Thus, CPI serves as a flexible and versatile integration platform, adapting to diverse enterprise needs across various SAP applications and beyond

3) Architecture

When we look into how CPI-DS is set up, it's kind of like other solutions in the CPI family, such as Smart Data Integration (SDI). Just like SDI, CPI-DS has two main parts. The first one is the Data Services Agent, which needs to be installed on-premise and handles communication on the source system side. The second part is the HANA Cloud Integration Agent Manager application, which acts as the server in the Cloud and is accessed by the (IT) user. Every so

often, the Agent checks with the CPI Cloud component to see if there are any queue jobs to process before taking a short break. Communication to and from the CPI Agent Manager, which you access through your web browser, happens securely over HTTPS. Below, a picture shows data flows with bold arrows and communication flows with regular ones. This setup helps make sense of the complex data and communication paths in CPI-DS.

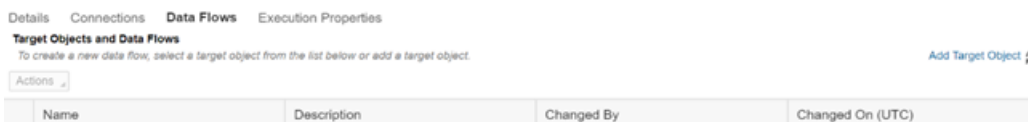


With CPI-DS, the main processing is handled by the Agent, while the Cloud component focuses on communicating job metadata based on user preferences – essentially, what needs to be done and how. It's vital to ensure that the Agent has enough resources, especially internal memory, to manage data volume peaks effectively. Since data loads cannot be split but only allocated to different agents, sufficient internal memory becomes crucial for smooth operations. Considering factors like data size, logic, and transformations, SAP recommends a minimum of 8 GB of internal memory for CPI-DS development/test environments and at least 16 GB for production environments with high data volumes. Having a second Agent deployed on another server, whether virtual or physical, not only provides fail-safe measures but also aids in load balancing. Agents can be grouped to optimize load allocation or configured to connect with other applications using protocols like SOAP, REST, and OData. It's important to note that installing both CPI-DS and SDI on the same

server isn't advisable, as they share essential resources like the database, which could lead to conflicts across multiple landscapes.

4) Core Function Overview

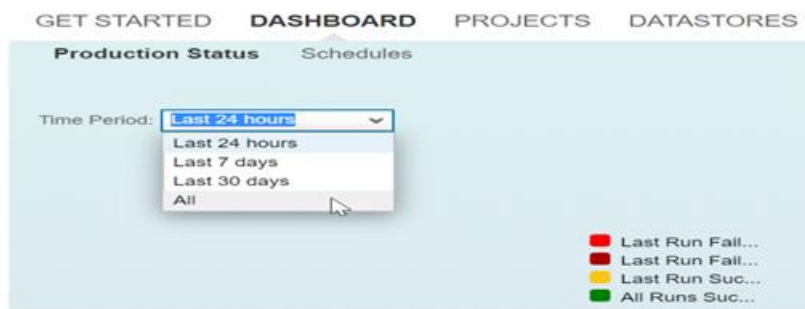
In CPI-DS, the key components related to data are tasks and data flows. A single task can encompass multiple data flows designed for Extract, Transform, and Load (ETL) processes toward specific targets. Tasks can be created either from scratch or using predefined templates. Data flows are crafted through the 'Task Editor' tab, which has a user interface similar to that of SAP BusinessObjects Data Services or the Eclipse Web IDE for HANA Native. However, the functionality of CPI-DS is comparatively limited in both cases. Each data flow must adhere to a standard format comprising initial transforms, intermediate transforms, and target transforms, with only the latter being mandatory.



When you're working with transforms, it's important to map the data columns appropriately. When setting flow targets, it's best practice to use the technical name of the table, and the file name should match the target in this scenario. You only need to map fields that are technically required, typically the key fields. Any fields that aren't mapped won't be altered by the transform process. If their respective columns are mapped, existing records will be overwritten. Specific properties like planning levels won't affect how this mechanism operates. Even if there's an incorrect planning level, the system will still overwrite a record with the correct planning level without triggering an error message.

Another handy feature in the data flow editor is the preview function. It's worth noting that this function is exclusively available in the Sandbox environment and is not accessible in the Production equivalent. The preview function lets you visualize the results of a specific query block within your data flow. The displayed results encompass the transforms up to that particular point in the flow, offering an opportunity to assess your flow's performance and alignment with expectations. Keep in mind that a maximum of 500 records can be displayed, and you can specify the system to make a selection per N-number of records. If this function isn't used, the system defaults to showing the first 500 records. Additionally, users can evaluate global data runs for either the

Sandbox or Production environments through the Dashboard work center, providing a historical perspective on the success of runs.



When considering user roles within the authorization concept, CPI-DS officially supports the same five roles as other CPI-suite solutions. However, in practice, user roles commonly revolve around Developers and Operators. Developers possess the capability to grant access rights to other users, manage connections and data flows, and configure data source registrations and tasks. On the other hand, Operators are tasked with executing tasks and flows, displaying and evaluating most objects, and conducting monitoring activities. SAP recommends assigning both roles to members of the IT department rather than business key users.

One final feature worth mentioning is agent monitoring in CPI-DS. While target systems typically have their monitoring functions for data loading processes, CPI-DS provides notifications for users to monitor the critical component - the Agent - in the loading process. The Administrator-work center enables users to set up simple yet valuable notifications for tasks, processes, and the Agent(s) used by CPI-DS. This limited automation helps address errors and provides support in scenarios where failsafe Agents may not be available within the organization.

Key Capabilities, Customer Benefits, and Considerations

- Gain direct, secure access to numerous SAP sources for seamless extraction, transformation, and loading into targeted cloud applications.
- Read and write data from diverse sources including databases (SAP HANA, DB2, Oracle, SQL Server) and files (XML, delimited).
- Utilize a role-based, wizard-driven web user interface in HTML5 to effortlessly design, execute, and monitor ETL (Extract, Transform, Load) jobs.
- Attain end-to-end visibility for data loads into IBP (Integrated Business Planning) and configure email notifications to stay informed about integration flow operations.
- Leverage the built-in scheduler or opt for the IBP scheduler. Additionally, integrate with third-party applications by invoking the integration flow through a web service call.
- Achieve seamless and secure integration from On-premise to SAP Cloud without the necessity of opening firewalls to inbound traffic.
- Data is transient in SAP HANA cloud integration, passing directly from source to target data-store within the SAP Cloud.
- Enjoy a guided user experience for effortless creation and editing of ETL (Extract, Transform, Load) jobs without the need for coding.
- Automate processes with "Set it and forget it" scheduling, eliminating manual intervention.
- Keep your on-premise footprint minimal (<100 Mb) with a software "agent" ensuring secure connectivity.
- Facilitates ABAP-based extraction and BW ODBC extractors for SAP Business Suite.

Volume 9 Issue 5, May 2020

www.ijsr.net

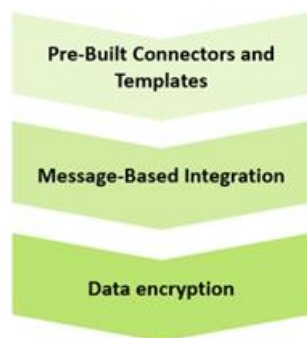
Licensed Under Creative Commons Attribution CC BY

- Streamline processes with automatic provisioning of sandbox and production repositories, enabling one-click promotion of jobs to production.
- Capable of supporting delta loads for efficient data management.

Advantages of SAP CPI –DS



Datastore serves as a vital link between CPI-DS/BODS and databases or applications, housing connection configurations. These configurations enable CPI DS/BODS to access metadata or perform read/write operations upon job completion. Notably, cloud-based applications exhibit better compatibility with SAP CPI DS compared to SAP BODS. CPI-DS is engineered as a cloud-native integration platform, harnessing the scalability, flexibility, and availability of cloud infrastructure. It seamlessly integrates with diverse cloud applications and services, simplifying data exchange between them. Offering built-in security measures, monitoring capabilities, and scalability to manage extensive data volumes, CPI-DS ensures smooth integration with other SAP Cloud Platform services and APIs.



CPI-DS offers a comprehensive array of pre-built connectors catering to both on-premise and cloud applications like SAP S/4HANA Cloud, SAP SuccessFactors, SAP Ariba, and more. These connectors streamline integration by furnishing ready-made connectivity, thereby reducing development efforts and time. CPI supports diverse messaging protocols and standards, facilitating event-driven and message-based integration. It enables the design and implementation of integrations responsive to events and messages, fostering event-driven architectures and near-real-time data processing. In contrast, BODS primarily serves as a batch-oriented data integration and ETL tool, though it does support near-real-time data processing, it's not explicitly designed for real-time and event-driven integration scenarios. CPI-DS offers data encryption features to safeguard against unauthorized access, providing message-level security. Asymmetric encryption, supported by CPI-DS, ensures a higher level of security compared to symmetric encryption methods. Private key storage on the CPI-DS server enhances security by storing the

private key in a file on the server, a more secure approach than storing it on the BODS server.

The user interface of CPI-DS boasts a user-friendly graphical interface (GUI), simplifying the creation and management of data integration projects. With CPI-DS, users benefit from a graphical environment where they can easily configure data loads through drag-and-drop functionality, offering a distinct advantage over SAP BODS.



CPI-DS offers a user-friendly experience, catering to individuals with limited data integration expertise. Its guided user interface assists in creating and managing data integration projects seamlessly. In contrast, BODS demands extensive technical knowledge to utilize its ETL capabilities effectively.

Furthermore, CPI-DS gains an edge through its cloud-based architecture, which enables it to leverage cloud infrastructure resources dynamically. It can adapt resource allocation based on demand, ensuring scalability requirements are met more effortlessly compared to an on-premises deployment of BODS.

In addition, CPI-DS empowers users to cleanse, transform, and enrich data before loading it into target systems, enhancing data quality and usability.

The platform also excels in monitoring and error handling, featuring sophisticated tools to track integration processes in real time. Detailed logs and monitoring capabilities enable users to identify bottlenecks or issues promptly and take necessary actions in case of failures or exceptions, ensuring data reliability and quality.

Database Comparison Chart

Database	CPI-DS	BODS
SAP HANA	Yes	Yes
SAP BW	Yes	Yes
Oracle	Yes	Yes
SQL Server	Yes	Yes
MySQL	Yes	Yes
CSV	Yes	Yes
XML	Yes	No
JSON	Yes	Yes
Web services	Yes	Yes
SAP MaxDB	No	Yes
Microsoft Access	No	Yes
IBM DB2	Yes	Yes
SAP Master Data Management (MDM)	No	Yes
SAP Lumira Cloud	Yes	No
SAP Cloud Platform HANA (SCP)	Yes	No
SAP Integrated Business Planning (SAP IBP)	Yes	No
Workforce Analytics Database	Yes	No

Other Benefits and Advantages

Extracting data from SAP S/4HANA to SAP Integrated Business Planning (SAP IBP) leverages SAP CPI-DS's capability to generate ABAP programs. This enables pushdown logic for heavy data processing, like complex joins from the DS-Agent Server to SAP S/4HANA Application Server, transforming data locally before transferring it to SAP IBP.

SAP IBP simplifies scheduling SAP CPI-DS jobs using the cross-system communication feature in the native Fiori app 'Application jobs'. Data integration is crucial for an end-to-end integrated business planning process. Explore available options for transferring mass data between SAP IBP and SAP S/4HANA, and remain open to new functionalities emerging in the market.

2. Conclusion

In this discussion we have included an overview of the SAP CPI architecture, which involves understanding how different components within SAP CPI interact to facilitate seamless integration between various systems and applications. This paper covers aspects such as message processing, connectors, runtime, and the overall structure of the integration platform. This paper covers how SAP CPI enables users to connect different systems, orchestrate processes, and manage the flow of data. Integration capabilities may include pre-built connectors, data mapping, content-based routing, and support for various protocols. CPI-DS provides a straightforward yet robust mechanism to facilitate batch and scheduled data integration, bridging the gap between diverse landscapes seamlessly. For scenarios where a more complex integration tool might seem excessive, CPI-DS emerges as a practical choice, offering efficiency and reliability without unnecessary complexity.

In summary, if you encounter integration hurdles within an SAP environment, especially concerning data exchanges between on-premises and cloud applications, it's crucial to give CPI-DS careful consideration. Capitalizing on its features enables you to streamline integration processes,

simplify complexity, and open doors to innovative data-driven solutions within your organization.

Declarations

Ethics approval and consent to participate: Not Applicable
Consent for publication: All authors have consent to submit this paper to the Journal of Cloud Computing. Also, we confirm that this paper or any part of this paper was not submitted anywhere.

Availability of data and materials: Not Applicable

Competing interests: Not Applicable

Funding: Not Applicable

Acknowledgments: Thank you co-author Maha Bhageshwara Raju Kesaboina for his expertise and assistance throughout all aspects of our study and for your help in covering a few topics and reviewing the manuscript.

References

- [1] "SAP Business Accelerator Hub," api.sap.com. <https://api.sap.com/integrations/packages>
- [2] Sookriti_Mishra, "My adventure in learning CPI: Part 1 | All about SAP Cloud.," SAP Community, Sep. 10, 2019. <https://community.sap.com/t5/technology-blogs-by-members/my-adventure-in-learning-cpi-part-1-all-about-sap-cloud/ba-p/13393526>
- [3] Sookriti_Mishra, "My adventure in learning CPI: Part 2 | Deployment Models," SAP Community, Sep. 15, 2019. <https://community.sap.com/t5/technology-blogs-by-members/my-adventure-in-learning-cpi-part-2-deployment-models/ba-p/13394902>
- [4] Sookriti_Mishra, "My adventure in learning CPI: Part 3 | Cloud Security," SAP Community, Sep. 15, 2019. <https://community.sap.com/t5/technology-blogs-by-members/my-adventure-in-learning-cpi-part-3-cloud-security/ba-p/13398951>
- [5] "SAP CPI | SAP Blogs," blogs.sap.com. <https://blogs.sap.com/tag/sap-cpi/>
- [6] STALANKI, "Comprehensive SAP CPI Guide for Standards & Best Practices," SAP Community, Jan. 16, 2020. <https://community.sap.com/t5/technology-blogs-by-members/comprehensive-sap-cpi-guide-for-standards-best-practices/ba-p/13457873>
- [7] "SAP Help Portal," help.sap.com. <https://help.sap.com/docs/cloud-integration/sap-cloud-integration/sap-cloud-integration>
- [8] "Home - Apache Camel," camel.apache.org. <https://camel.apache.org/>
- [9] "SAP Help Portal," help.sap.com. <https://help.sap.com/docs/cloud-integration/sap-cloud-integration/integration-capabilities>
- [10] engswee, "CPI's Groovy meets Spock – To boldly test where none has tested before," SAP Community, Jan. 24, 2018. <https://community.sap.com/t5/technology-blogs-by-members/cpi-s-groovy-meets-spock-to-boldly-test-where-none-has-tested-before/ba-p/13367099>

- [11] “Steps to learn SAP CPI, what you must know before your first project | SAP Blogs,” blogs.sap.com.
<https://blogs.sap.com/2019/12/06/steps-to-learn-sap-cpi-what-you-must-know-before-your-first-project/>
- [12] “The Apache Groovy programming language - Syntax,” groovy-lang.org. https://groovy-lang.org/syntax.html#_unicode_escape_sequence
- [13] “Comprehensive SAP CPI Guide for Standards & Best Practices | SAP Blogs,” blogs.sap.com.
<https://blogs.sap.com/2020/01/16/comprehensive-sap-cpi-development-standards-best-practices/>
- [14] “Integration Developer’s Corner | SAP Blogs,” blogs.sap.com.
<https://blogs.sap.com/2017/07/17/introduction-to-developers-corner/>
- [15] “SAP Help Portal,” help.sap.com.
<https://help.sap.com/docs/cloud-integration/sap-cloud-integration/integration-flow-design-guidelines>
- [16] “Externalizing parameters using SAP Cloud Platform Integration’s Web Application | SAP Blogs,” blogs.sap.com.
<https://blogs.sap.com/2017/06/20/externalizing-parameters-using-sap-cloud-platform-integrations-web-application/>
- [17] “Enrichments of Externalization Feature in SAP Cloud Integration | SAP Blogs,” blogs.sap.com.
<https://blogs.sap.com/2020/01/21/enrichments-of-externalization-feature-in-sap-cloud-platform-integration/>
- [18] “SAP Cloud Integration for data services,” support.sap.com.
<https://support.sap.com/en/alm/solution-manager/expert-portal/public-cloud-operations/sap-cloud-platform-integration-for-data-services.html>