Employee Central Data Migration: Cutover Optimization Strategy Using Infoporter
The recommendations in this document are based on the functionality available up to SAP SuccessFactors release mentioned above. Future functionality can impact the recommendations provided by this document. We strive to keep these recommendations up-to-date, however, in case you find that recent new functionality has not yet been considered in the latest version of this document, please reach out to your Customer Success Manager / Partner Delivery Manager or send an email to SAPSuccessFactorsIDPDoc@sap.com.

Implementation Design Principles (IDPs) for SuccessFactors solutions are delivered by SAP for helping customers and partners on how to choose the most appropriate strategy and solution architecture for SuccessFactors implementations. IDPs are compiled taking into consideration the experience of many implementation projects and addressing frequent business requirements as well as real-life implementation challenges. They are continuously reviewed and updated as product functionality evolves. In addition, the reader is advised to read and familiarize with essential and additional product-related documentation which includes Implementation Guides, SAP Notes, SAP Knowledge Base Articles, and additional assets as referenced in this document, see chapter 6.
TABLE OF CONTENTS

1 TERMINOLOGY ........................................................................................................... 4

2 ABSTRACT .................................................................................................................. 4

3 INTRODUCTION ......................................................................................................... 5

4 BUSINESS REQUIREMENT ......................................................................................... 5

5 DETAILED SOLUTION ................................................................................................. 5

5.1 INFOPORTER FUNCTIONALITY ............................................................................. 5

5.2 CSV VS WEB SERVICES ......................................................................................... 6

5.2.1 Recommendation ............................................................................................... 6

5.3 CUTOVER PROCESS FOR DATA MIGRATION AND REPLICA CTION .................... 7

5.3.1 General cutover process and transactional freeze ............................................. 7

5.4 CUTOVER PROCESS VARIANTS DEPENDING ON MIGRATION SCENARIO ........... 9

5.4.1 Replication to the same SAP ERP HCM system .............................................. 9

5.4.2 Data Migration from different SAP ERP HCM system(s) .................................. 9

5.4.3 Replication to new SAP ERP HCM system(s) ................................................. 9

5.4.4 Data Migration from SAP ERP HCM to EC .................................................. 10

5.5 OPTIMIZING THE TRANSACTION FREEZE PERIOD ........................................... 10

5.5.1 Cutover testing ................................................................................................... 10

5.5.2 Data Quality Impact .......................................................................................... 10

5.5.3 Cutover Preparation Work ................................................................................ 10

5.5.4 Data Migration Initial Load .............................................................................. 11

5.5.5 Cutover Plan ..................................................................................................... 11

5.5.6 Point of No-return ............................................................................................ 12

5.5.7 Operation Modes .............................................................................................. 12

5.5.8 Employee Central Rules .................................................................................. 12

5.6 OPTIMIZING DATA MIGRATION WITH WEB SERVICES ................................. 13

5.6.1 Using the Job Schedulers to run Data Migration Jobs ..................................... 13

5.6.2 OM Data Migration .......................................................................................... 14

5.6.3 Position Data Migration ................................................................................... 14

5.6.4 Employee Data Migration ............................................................................... 14

5.6.5 Disable HRIS Sync ........................................................................................... 15

5.6.6 Testing Employee replication in simulation mode .......................................... 15

5.7 OPTIMIZING DATA MIGRATION WITH CSV ................................................... 15

5.7.1 Performance Settings ....................................................................................... 15

5.7.2 Splitting the Import File ................................................................................... 16

5.7.3 Organizational Data Migration ........................................................................ 16

5.7.4 Position Data Migration .................................................................................. 16

5.7.5 Employee Data Migration ............................................................................... 17

5.8 OPTIMIZING INITIAL REPLICA CTION FROM EC TO SAP ERP HCM .......... 17

5.8.1 Tuning SAP ERP foreground and background process types ......................... 17

5.8.2 Replication process depending on object type ............................................... 17

5.9 PERFORMANCE OPTIMIZATION FOR EMPLOYEE DATA REPLICA CTION ......... 19

5.9.1 Process settings ............................................................................................... 19

5.9.2 Data Replication Monitor ................................................................................ 20

6 REFERENCES ............................................................................................................. 21
1 TERMINOLOGY

The following table explains some abbreviations used in this document.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAP ERP HCM</td>
<td>The SAP ERP HCM system</td>
</tr>
<tr>
<td>EC</td>
<td>The SuccessFactors Employee Central System</td>
</tr>
<tr>
<td>SOR</td>
<td>System Of Record. The Single Source of truth that feeds master data to other systems.</td>
</tr>
<tr>
<td>Cutover</td>
<td>“Cutover” is a general IT project management term used to refer to the activities which have to be performed for switching from an old to a new system. In the context of Employee Central implementation projects, the “cutover” term usually encompasses a broad list of activities starting with setting up the productive instance’s provisioning settings, loading data model XMLs, loading diverse configuration into the productive instance, connecting interfaces, loading certificates, and many such activities. The actual masterdata (employee and organizational data) cutover activities come at the later point of the overall cutover process. This document focuses solely on the aspects regarding “masterdata cutover activities”. For the sake of simplicity, in this document these are simply referred to as “cutover activities” or “cutover phase” or simply “cutover”.</td>
</tr>
<tr>
<td>Transactional Freeze Period</td>
<td>The term we use to describe the cutover period in which data on original production systems is frozen, data load into Employee Central is performed and initial replication from EC to the original systems take place.</td>
</tr>
<tr>
<td>Migration</td>
<td>The one-time process of moving data in bulk from one system to the other. Typically, from SAP ERP HCM to EC. Sometimes also called Replication from SAP ERP HCM to Employee Central</td>
</tr>
<tr>
<td>Initial Replication</td>
<td>The one-time process of moving data in bulk from Employee Central to SAP ERP HCM.</td>
</tr>
<tr>
<td>Replication</td>
<td>The continuous process of sending changes from the SOR system to downstream systems. In this case, it is used in the context of Replicating changes from EC to SAP ERP HCM after the cutover.</td>
</tr>
<tr>
<td>Infoporter</td>
<td>The tool delivered by SAP and installed in the SAP ERP HCM system as part of the EC licenses to migrate and replicate data between SAP and EC.</td>
</tr>
<tr>
<td>WS</td>
<td>Infoporter Web Services. The method whereby you use the web services functionality of the Infoporter to move or replicate data in a one step process.</td>
</tr>
<tr>
<td>CSV</td>
<td>Infoporter CSV files. The method whereby you use the Infoporter to generate flat files (CSV) to extract data from SAP ERP HCM as CSV files and upload them into EC in a two-step process.</td>
</tr>
<tr>
<td>BIB</td>
<td>Business Integration Builder. The SAP integration tool, accessed through the IMG, to configure the Infoporter Templates in a template group.</td>
</tr>
<tr>
<td>Template</td>
<td>The Template contains the configuration and mapping for replicating data to and from an Employee Central Entity.</td>
</tr>
<tr>
<td>EC Entity</td>
<td>The name for a storage location where data is stored in Employee Central. The content of an EC Entity is accessed through an EC portlet or by using manage data for an MDF object.</td>
</tr>
<tr>
<td>CPI</td>
<td>Cloud Platform Integration. The middleware used by the Infoporter to migrate data using web services to EC, and to replicate data from EC to SAP ERP HCM.</td>
</tr>
</tbody>
</table>

2 ABSTRACT

Cutover processes are crucial in any cloud migration project. Especially during cutover time-window starting with the initial data load to the new system until releasing it for productive usage. To reduce production downtimes and ensure smooth business continuity, it is important to keep this specific cutover time-window as short
as possible. This IDP provides ways of optimizing the data migration and replication cutover process for Employee Central implementations with the aid of SAP Infoporter solution. It discusses cutover activities like delta migration, freeze period, testing data replication and performing the initial replication from Employee Central.

3 INTRODUCTION

“Cutover” is a general IT project management term used to refer to the activities which have to be performed for switching from and old to a new system. In the context of Employee Central implementation projects, the “cutover” term usually encompasses a broad list of activities starting with setting up the productive instance’s provisioning settings, loading data model XMLs, loading diverse configuration into the productive instance, connecting interfaces, loading certificates, and etc. The actual masterdata (employee and organizational data) cutover activities come at the later point of the overall cutover process. This document focuses solely in the aspects regarding “masterdata cutover activities”. For the sake of simplicity, in this document these are simply referred to as “cutover activities” or “cutover phase” or simply “cutover”.

4 BUSINESS REQUIREMENT

During a specific time-window of the Employee Central implementation cutover, a freeze of the HR data will be required in productive systems, leaving some HR processes temporarily unavailable. Therefore, it is important to keep this time-window as short as possible. This document describes how to optimize Employee Central cutover activities related to data migration and initial replication in order to achieve a shorter data freeze time-window.

This document assumes an Employee Central Core Hybrid deployment in which the migration of data from SAP ERP HCM to Employee Central is done using the Infoporter tool. Note that cutover for Employee Central Payroll (ECP) is not in scope of this document, as Infoporter cannot be used for ECP migrations.

5 DETAILED SOLUTION

5.1 Infoporter Functionality

The Infoporter functionality has already been described in detail in Implementation guides. The following is a summary of Infoporter aspects relevant to our discussion.

Infoporter is the SAP recommended technology for migrating employee masterdata and organizational data from SAP ERP HCM to EC. It is installed on the SAP ERP HCM system as an Add-On with software component PA_SE_IN. It contains built-in logic for extracting data with HR processes such as Global Assignment, Hire, Retire, Termination. It is configured using the Business Integration Builder (BIB) framework. This framework is used for both data migration (Infoporter’s scope) and data replication (from EC to SAP ERP HCM). The Infoporter can only be used on SAP systems. If you are not migrating from an SAP ERP HCM system Infoporter cannot be used.

Within the SAP Infoporter there are two data transfer methods: Web services transfer and CSV file transfer.

<table>
<thead>
<tr>
<th>Data type / activity</th>
<th>From SAP to EC</th>
<th>From EC to SAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foundation data</td>
<td>Web services since 2019</td>
<td>Not possible</td>
</tr>
<tr>
<td></td>
<td>CSV is used but not generated by the Infoporter</td>
<td></td>
</tr>
<tr>
<td>Data Migration</td>
<td>Web Services or CSV</td>
<td>Web Services</td>
</tr>
<tr>
<td>Initial Replication</td>
<td>N.A. (see Data Migration)</td>
<td>Web Services</td>
</tr>
<tr>
<td>Continuous Replication</td>
<td>Web Services</td>
<td>Web Services</td>
</tr>
</tbody>
</table>
5.2 CSV vs Web Services

Choosing whether to use CSV or Web Services is an important decision that needs to be taken at the beginning of the project, as it leads to two different approaches. There is no one size fits all recommendation. Both options have a cost impact. Once you have chosen a method, the relevant EC entities metadata needs to be uploaded to the BIB, and switching technologies means that you need to restart the configuration. You cannot simply convert CSV Templates to WS Templates and vice-versa.

You can take the following into consideration when choosing:

Scope:
- CSV can only be used for Data Migration from SAP to EC. For replication, web services are the only option.
- If you migrate from SAP to EC, and then need to replicate changes back, you can reuse the Data Migration WebService configuration for the Replication back as it is the same tool and uses the same templates. These may have to be copied to a new template group if there are differences. You cannot use Data Migration CSV templates for Replication.
- Note that prior to version 1611 only CSVs could be used for Data Migration, and if your project started then it was the only option.

Migration Effort:
- Web Services scale easily and can run in parallel jobs. This splitting is built into the Infoporter migration programs and you can run 10 to 50 jobs in parallel, all handled by the Infoporter and all logs are centralized.
- When using CSV it’s recommended to split files into sets of 10,000 records per portlet for performance reasons, though the hard limit is 50,000. This splitting is a manual process, which can take up excess project time and needs to be repeated multiple times. If you have 10,000 Employees and 20 EC Entities, you will have 20 files. If you have 30,000 employees and 20 EC Entities, you will need to split manually and have 60 files. Only 5 can be uploaded in parallel and there will be an error log file per uploaded file.

Volume of changes and cutover window
- If the cutover window is small, Web Services are faster, and you can process delta loads using the same technology. Note that CSV based file transfer does not support delta migration.
- The Infoporter can do an initial load with web services, and then, by enabling the Delta Replication in the Template group configuration, use standard SAP change pointers to track all changes since the last upload and migrate the changes since the last run. This delays the start of the data freeze.
- With CSV, the cutover window starts when you do the first extract, as there is no straightforward delta functionality.

Configuration effort:
- Web Services require more configuration work. As there are no files created all mapping needs to be done in the Infoporter configuration. The benefit is that it is consistent and automated.
- CSV require the same effort, but there may be cases where the decision is taken to manually update the CSV content for specific cases which require too much effort to automate. This manual update needs to be performed every time you extract and reupload the data during testing. The manual changes need to be documented, followed, repeated and is prone to mistakes (forget a step, not doing the same step in the same way).

An example can be the address format which is different, and it’s decided that for smaller countries it’s updated manually. This may require localized knowledge and possibly the data is not in a language with the migration specialist understands.

Sometimes there may be a mix, where global data is automated using web services, and specific exotic cases are handled using CSV. This can be an option for small volume records.
Accuracy, Security and confidentiality:

- **Webservices**:
  - All modifications / mapping to the data are done in the configuration and transported to production using transports.
  - All data is transmitted to EC over encrypted https Web Services
  - No files are generated, it is a one step process.
  - No manual splitting is required.
  - No manual modifications are possible to the data

- **CSV**
  - Modifications can be done to the files after extracting from SAP ERP HCM and before importing to EC. These may or may not be documented and user errors can occur.
  - A copy of the data resides on file servers for a while and needs to be cleaned up.
  - Manual splitting may be required.

Error Analysis

- **Web Services**
  - Executing the web service call from SAP ERP HCM will get an immediate success / error message with details if there is an error

- **CSV**
  - Executing the export from SAP ERP HCM will not get any immediate import error messages.
  - The error will be reported during the next step during and after the upload to EC, which may be executed by a different person

The error may also be caused by format error, local pc settings, splitting or manual changes to the content.

5.2.1 **Recommendation**

Web Services are recommended for any scenario where you have:

- Multiple countries (+ 10) and more than 10.000 Employees.
- Single country as of +25.000 employees

Note that once you made a choice, the template configuration, metadata and fieldnames are different for CSV and Web Service templates. There is no automatic conversion from one type to the other and you cannot easily switch between the two methods.

Web Services Summary:

- More configuration effort is required, which will be offset by less effort during the data migration test and live runs.
- If you also need to replicate back from EC to SAP ERP HCM, you can reuse that configuration.

CSV Summary:

- Less configuration effort, especially if replication is not in scope
- More migration effort and a longer Transactional Freeze Period for cutover.

Note that there are large scale projects that did use CSV. It is possible, but it requires a longer cutover window.

This was either due to:

- A lack of understanding of the web services benefits and/or the customer decide to manage the data migration using CSV.
- The project scope and approach were defined before enhanced web services were released for data migration (v 1611).

5.3 **Cutover Process For Data Migration and Replication**

This chapter explains the most common integration scenarios and the different cutover phases involved.
5.3.1 General cutover process and transactional freeze

The Employee Central cutover process is a period of multiple weeks containing activities which can be high-level described as follows:

- **Preparation work**
  - Deploy the EC system
  - Configure the EC system
  - Load Foundation data
    - Cost centers, Picklists, etc.
  - Load other data relevant for the functional configuration of the EC system
- **Start of data freeze period in the source system to migrate transactional data** (Time sensitive period)
  - Data Migration
    - Load actual OM and PA data
      - Org structure
      - Positions
      - PA data
  - Validation of the results of the Data Migration
  - Initial Replication
    - Align EC data with SAP data
  - Tasks that are required before the EC system is released for changes
- **End of Transactional Freeze Period**
  - Data changes can be performed in EC by a specific pilot user group
- **Preparing EC for general live use**
  - Activate rules, permissions and other such activities …
  - Activate replication of changes performed in EC to SAP.
  - System is live.

The data freeze period explained above can be further detailed into two phases:
- a limited transactional activity phase, and
- a hard-transactional freeze phase.

The limited transactional activity phase is made possible due to the “delta replication mode” offered by the Infoporter solution when using Web Services. During this phase, urgent and business-required changes can still be performed in the original SAP ERP HCM productive system(s). The real “hard data freeze” period is then further reduced with this approach. The figure below shows an example of how the data migration and replication cutover process may look like:
5.4 Cutover process variants depending on migration scenario

5.4.1 Replication to the same SAP ERP HCM system

This is the common scenario that is challenging for cutover planning as the data that needs to be loaded from SAP will be updated by the replication from EC as part of the cutover.

A single SAP source System migrates to a single EC target system and replicates back to the same SAP system. Within the Transactional Freeze Period you need to:
• Migrate data from SAP to EC
• Perform initial replication from EC to SAP

Typically, all these steps need to happen consecutively and fast to ensure a short Transactional Freeze Period. You cannot descope something and need to optimize the steps.
As the initial replication from EC can overwrite the data in SAP, which is your source system, all relevant data must be correctly migrated to ensure no data loss.

5.4.2 Data Migration from different SAP ERP HCM system(s)

This scenario is less common, and potentially more complex. Historically this was the original positioning of Employee Central that acts as a centralized hub to merge data from multiple SAP HR systems.
Within the Transactional Freeze Period you need to:
• Migrate data from multiple SAP systems to EC
• Perform initial replication from SFEC to multiple SAP systems

It may be possible that the mapping for Data Migration and Replication is different, and that you cannot reuse the Infoporter configuration. In this case, it is very common to go live in waves for the PA data, i.e. perform a separate cutover for each SAP system. This depends on the situation, and the waves may follow one week after the other, or there may be weeks or months between the waves. Note that for OM and Positions you should not go live in waves. It’s recommended to have, always a single system of Record where you perform the changes to the Org structure / Position structure.

5.4.3 Replication to new SAP ERP HCM system(s)

This scenario is less common, and more complex: a single SAP source System migrates to a single EC target system but replicates back to one or more different SAP systems and not to the original SAP source system.
Within the Transactional Freeze Period you need to:
• Migrate data from SAP to EC
• Perform initial replication from EC to one or more different SAP systems
This happens when you have a Global SAP Master system, feeding other SAP HR systems and you replace the SAP Master System with EC while retaining the other SAP (HR) systems. This may be planned to happen in waves which will optimize your cutover, or all systems together. In theory you would not need all data to replicate to EC and each system needs their own subset of data. Furthermore, it's very common that these systems, as they were split before, do not need continuous updates, and this can be handled by different teams on different schedules. In some cases, systems only need updates on a weekly basis. Note also that as there is no Initial Replication to the SAP source system, the data in SAP stays as is, and can be consulted in case of issues. It may also be possible that the mapping for Data Migration and Replication is different, and that you cannot reuse the Infoporter configuration for each system.

5.4.4 Data Migration from SAP ERP HCM to EC

This is the least complicated scenario as you do not need to replicate back to SAP: a single SAP source System migrates to a single EC target system and there is no replication back. It's a onetime move and no integration between the two system is required. You stop using SAP ERP HCM.

Within the Transactional Freeze Period you need to:
- Migrate data from SAP ERP HCM to EC

Your Transactional Freeze Period starts when you start the data extraction and finishes when the data has been successfully loaded into EC and the critical tasks have been performed. In this scenario, you move from SAP ERP HCM to EC and stop using SAP ERP HCM. Note also that as there is no Initial Replication to the SAP ERP HCM system, the data in SAP stays as is, and can be consulted in case of issues.

5.5 Optimizing the Transaction Freeze Period

You optimize the Transactional Freeze Period by:
- preparing, testing and based upon your results anticipating all possible activities to the Cutover preparation phase before the Transactional Freeze Period, thereby minimizing the actual activities to be performed during the Transactional Freeze.
- Moving non-time-critical tasks to the post transactional freeze period.

Some activities may need to be part of the Transactional Freeze Period. For example, activating role-based permissions which first require the data to be loaded in EC. In any case, many activities can be pushed to post-freeze period, such as running the HRIS-Sync to populate the Employee Profile (as the replication does not read data from the Employee Profile).

5.5.1 Cutover testing

The recommended way to properly plan a cutover is by testing it at least once with the complete scope, using the same steps you will use for the go-live, and at the same speed. As you perform these steps they need to be documented in detail with parameters used and dependencies and with the duration and runtime. Even if you have errors during testing, you can still use the runtimes and duration to plan your actual cutover, provided the error volume is low. You may need to perform multiple tests of the scope to ensure data correctness and carryout functional testing to see end to end migration/replication works as desired before the actual cutover.

5.5.2 Data Quality Impact

When there are issues with data quality and you get errors, the performance test results may be less consistent as errors impact the performance and therefore the duration. This is mostly a problem during tests.
as you should not go live and cutover if you have data errors. The assumption is that these have been solved during the Cutover testing.

- **During Data Migration** the load performance will vary: If the user creation fails, the other templates will fail, and this will go faster than if they would really load.
- **If the user creation succeeds**, but you have other errors the data load will take longer due to the processing of the error message.

During Replication, the process usually is faster as these records will be skipped when you have errors.

### 5.5.3 Cutover Preparation Work

Many activities can be done in advance before you start moving the actual data. Anything that can be done before the data freeze leads to less impact. You can setup and prepare the EC system:

- Create the users
- Load the foundation structure
- Load all other configuration

This can be done using uploads or by using Instance Sync. This data does not change a lot (example: Locations do not change, and customers seldom perform a cutover during a company reorganization or restructuring)

Which option you choose is less important, however, you must use the same option both in the final test and the actual go live to ensure you can accurately document and benchmark the cutover process.

The only exception would be if you can go-live in production using a method that was not tested in the specific customer environment. It’s very unlikely that the implementing customer does not have successful testing as a required step in their audit process before moving new software into production.

### 5.5.4 Data Migration Initial Load

The Data Migration initial load is the biggest factor that can reduce your Transactional Freeze Period.

If you are using the Infoporter and are using Web Services, you can do an initial load of all OM and PA data from SAP ERP HCM into SF EC before the actual Transactional Freeze Period and then use the “Delta” functionality or simply perform a full reload which is highly automated when using web services.

An initial load with CSV is possible, but it does not support delta loads. Note that you need to disable HRIS Sync during this initial load, else the Employee Profile will be updated which will impact other SF modules if they are in use. As part of the final activities of the cutover, HRIS Sync needs to be reenabled again.

To activate Delta processing for Web Services you activate ‘Enable Delta Replication’ on template group level before your last full load of data. At the moment of activation, change pointers will be generated for each changed record in scope of the template group which you can then process after the last full load and before the go live. You can also process delta changes on a recurring basis before go-live to ensure that during the data freeze you only have a small set of delta changes to process. Note that you can only activate change pointers for one template group. So, all templates in scope should be in the same template group. Templates in other groups will not be able to use the delta replication.

As after your initial load you can then process the change pointers and process the changes between your initial load and your final load, this also means that the Data Freeze starts before your last Delta load, which can be weeks after your initial load. In theory more than 95% of your data is already loaded before your Transactional Freeze Period starts and only small delta loads are required thereafter to finalize the complete data migration.

If for some reason delta loads are not possible, a full reload of the data is also an option: delta loads cannot be spread across different jobs automatically (cannot be parallelized). Therefore, in some cases if enough background processes are available on the SAP Application server, a spread of 30 full load jobs will be faster than a single delta load job.
5.5.5 Cutover Plan

Cutover-plan templates are not provided as part of this IDP as it is task of the implementation partner to compose an appropriate project-specific cutover plan. A cutover plan is a sequential list of steps that need to be performed for the go-live. You build this plan as you test the cutover process and document every individual step, validating whether the step can be performed before, during or after the Transactional Freeze Period. Combining the plan with the actual results of the runtime will give you the time required for the Transactional Freeze Period and how to plan for it. The plan will also allow you to understand duration of the activities, and where it may be useful to optimize.

5.5.6 Point of No-return

The point of no (easy) return is the moment from which if you proceed, the actions to abort go-live and reverse the steps are disruptive to the business. When planning a cutover, you need to determine the point of no return. This will be different in every scenario. For example, if you migrate to a new EC system and no other modules are in scope, the point of no return could be:

- When replication starts updating SAP ERP HCM with the data from EC (if replication is in scope)
- When processes or people start updating data in EC and no longer in SAP ERP HCM (you start using EC).

Before that moment, you can still stop the cutover without the need for a restore of the system. If other modules are in scope and were already live, the point of no return could be the moment where EC HRIS Sync starts populating the Employee Profile, replacing the existing Employee Profile which used to be populated by SAP ERP HCM.

It’s important to understand what the actions are that need to take place to cancel the cutover.

- Before the point of no return: For example, announce the cutover is delayed and continue using SAP ERP HCM.
- After the point of no return: For example, announce the cutover failed and SAP ERP HCM needs to be restored.

5.5.7 Operation Modes

It’s recommended to use predefined Operation Modes on SAP ERP HCM to easily switch between data migration mode and replication mode. This is typically setup by the SAP Basis group to not interfere with their Operation Modes.

- Data Migration requires many background processes
- Replication requires a lot of dialog processes

As it’s not viable to restart the SAP ERP HCM system for this, a simple solution is to ask your SAP Basis team to create Operation Modes for the same system, with a migration operation mode with many background processes, and a replication mode with many dialog processes. The operation mode switch will change the type of process that can be used, from the total pool of available processes. SAP Basis should be aware of the OP mode functionality, and it’s probably already in use. This is an SAP standard functionality.

Note that if you are using CSV, while it also uses background processes, it needs to generate files. So only the background processes of the specific application server where you generate the files is relevant, and not the total pool of available background processes.

5.5.8 Employee Central Rules

Quite frequently EC business rules are implemented to update Job Info when changing positions and vice versa. This can triple the duration of data loads (based on observations). You can test this by doing data loads with the rules on and off. A good option is to disable these rules, load the correct data into the correct portlet as part of the data migration and only then activate the rules.
5.6 Optimizing Data Migration with Web Services

Data Migration with Infoporter via Web services is an SAP ERP system background process. This means you need to have background processes available on your SAP instance (see Operation Modes). When running the Data Migration programs, OM Data Migration works different from PA Data Migration, and different combinations of number of parallel jobs and objects per job will be required.

Depending on the number of EC Entities you send data to, more jobs can run in parallel. However, if you run too many jobs in parallel you may get performance related errors, which also takes time to resolve. You determine this optimal load pattern during the testing.

5.6.1 Using the Job Schedulers to run Data Migration Jobs

Both for OM and PA data there are programs to run the data migration and programs that can schedule multiple instances of the data migration programs. These are job schedulers for large volume datasets.

For OM, the Job scheduler in SAP ERP HCM is report ECPAO_OM_OBJ_DMT_JOB_SCHEDULER. It will schedule multiple instances of ECPAO_OM_OBJECT_EXTRACTION. As there are less EC “entities” related to OM objects, the default settings for number of objects per job and number of parallel jobs is lower than for PA jobs. For PA, the job scheduler report ECPAO_EMPL_DMT_JOB_SCHEDULER will schedule multiple instances of ECPAO_EMPL_EXTRACTION. You provide to the schedulers the Object list, the number of jobs in parallel, and the number of objects in a single job.

The total number of jobs will be the total number of objects divided by the number of objects per job and is determined by the job scheduler. The job schedulers by default proposes the number of objects and jobs in parallel and will give a warning when you increase that value. The defaults are the recommend values. You can change those values but because you are using Cloud Platform Integration (CPI) as the middleware, it’s important to understand that the CPI will split the load into smaller batches of 100 objects in any case.

Sending more than 100 objects per job will increase the run time as CPI needs to do further splitting. You can speed up that process by sending jobs of 100 objects or less.

Examples and recommendations:
Note that this is mostly theoretical, as typically you will be moving +10,000 records, and will have more than 100 jobs.

1,200 objects with 10 jobs in parallel and 100 objects per job.
- This will result in 12 jobs.
- 10 jobs will run, and the next 2 jobs will start as and when the previous job finished.

1,200 objects with 10 jobs and 120 objects per job.
- This will result in 10 jobs.
- These jobs will take more time, but you may end up being faster because you do not have 2 jobs waiting to start until 2 of the 10 existing jobs are finished.

It is recommended to split volumes into small batch sizes and using the default settings in the respective programs. This reduces the impact of a single failed job. The defaults for each program have been chosen based upon this.
If you want to improve the performance, you should change the quantity of jobs in parallel, and not the objects per job unless you are loading a very small quantity (example: 50 objects with the default of 100 objects will only result in 1 job, you can use 5 jobs with 10 objects each). Note that 100 jobs in total with 50 objects each is usually faster than 10 jobs with 500 objects. The program will give a warning when you increase these values. During projects, it is ok to perform more than 2,000 short jobs per day during the data migration phase.

Migrating 50,000 Employees and their respective positions using 100 objects per job will result in:
- 500 jobs for the positions without relationships
• 500 jobs for the positions with relationships
• 500 jobs for the PA templates without relationships
• 500 jobs for the PA templates with relationships

5.6.2 OM Data Migration

The quantity of Organizational objects (Type O) is typically low, and their relationships do not change frequently. It may be easier to do a simple full reload multiple times, rather than to try to setup delta processing for OM. The OM Delta processing is limited and may not capture all relevant changes. If you change an org unit, those changes will be captured, but as a lot of data is inherited, underlying org units are not flagged as changed because there is no change and they will not receive a change pointer. Full loads usually take less than 30 minutes for less than 10,000 objects using the default of 10 jobs and 50 objects. Therefore, consider performing only full loads into EC for OM objects.

5.6.3 Position Data Migration

When migrating Positions, all data goes to a single Position MDF Object. Running too many migration jobs in parallel for position objects will cause an overload on the Position MDF objects and this will result in Server errors.

You optimize the position objects migration load by testing out different combinations of parallel jobs and objects per job until you get performance bottleneck errors. The default number of parallel jobs and objects per job for the OM program for positions is a good choice. It's best to start with the default number of parallel jobs and objects per job, and only increase or decrease the number of parallel jobs. You can then find the performance bottleneck by monitoring the performance errors, and then reduce slightly to find the optimal distribution. As a rule, it will be challenging to have more than 10 jobs in parallel loading data into a single EC Entity and it is recommend using no more than 10 parallel jobs with 50 objects per job for positions.

5.6.4 Employee Data Migration

When migrating PA data, the data goes to different EC Entities. This means the performance bottlenecks will be different, as any load will process more than 1 template, and therefore go to more than one EC Entity. Therefore, it is possible to use a higher number of jobs in parallel.

The default is 10 jobs with 100 objects per job. This works for most systems. If you are loading data for more than two EC Entities, you could add 5 parallel jobs at a time to find performance bottlenecks by looking for Soap or server error messages. A standard SAP ERP HCM and EC system should normally not have many issues with running 10 jobs in parallel. If you need more due to volumes, it is assumed that your system is performant enough, as it has to cater for more employee data. You will also need free Batch processes to be able to run this migration.

With 10 jobs having 100 objects per job, it is typically possible to load 5,000 to 10,000 PERNRs per hour. More may be possible, but this is a good starting point. Similarly, as you have more data, you will be able to use more jobs. As of 30 jobs in parallel you may encounter performance bottlenecks if components are not properly sized.

You can use the following options and expect the following results, assuming 20 EC Entities. These are not official benchmarks, based upon observation on some actual projects.

<table>
<thead>
<tr>
<th>Total nr PERNR</th>
<th># Parallel Jobs (2)</th>
<th>Objects per job</th>
<th>Pernr / Hour (1)</th>
<th>Total duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20,000</td>
<td>10</td>
<td>100</td>
<td>10,000</td>
<td>2 h</td>
</tr>
<tr>
<td>20,000 – 50,000</td>
<td>20</td>
<td>100</td>
<td>10,000 – 15,000</td>
<td>4 h</td>
</tr>
<tr>
<td>50,000 – 100,000</td>
<td>30</td>
<td>100</td>
<td>10,000 – 20,000</td>
<td>4 h</td>
</tr>
<tr>
<td>+100,000</td>
<td>40 – 60</td>
<td>100</td>
<td>12,500 – 17,500</td>
<td>&lt; 8 h / 100,000</td>
</tr>
</tbody>
</table>
(1) Note that the performance may go down a lot if ABAP / BADI coding is used to transform or lookup data while executing the templates.

(2) This is for Data Migration with WS, note that with CSV the maximum is 5 jobs in parallel

### 5.6.5 Disable HRIS Sync

When migrating employee data using Web Services, the Odata API will execute the HRIS Sync. This will slow down the system. As you are splitting data migration jobs into multiple loads, it’s better to disable HRIS sync during the loading, and enable it just before go-live.

See this note which explains what HRIS Sync is. OSS Note 2080728
https://launchpad.support.sap.com/#/notes/2080728

To disable it, you can schedule a single, empty HRIS sync job and delete all others. This requires provisioning access.
See OSS note 2349390
https://launchpad.support.sap.com/#/notes/2349390

After the data load you will need to run a full HRIS Sync. This requires provisioning access.
See OSS note 2263251
https://launchpad.support.sap.com/#/notes/2263251

### 5.6.6 Testing Employee replication in simulation mode

You should not run actual employee data replication (from EC to SAP ERP HCM) until all data has been migrated successfully. You can however use the simulation mode (in the production environment). As you have done an initial migration from SAP ERP HCM to EC with Web Services, and then processed deltas, you already have data in EC and can test your masterdata replication from EC to SAP ERP HCM in simulation mode. You can then look at the replication log and check for errors. This means you can test replication and performance before the actual Transactional Freeze Period.

### 5.7 Optimizing Data Migration with CSV

Migration with CSV is not recommended for large volumes, as it is slower, and you have less flexibility to spread the load. It is however possible. The process is mature and there is benchmark data available that can be used as a reference.

As there is no Delta functionality, you must execute the complete process during the Transactional Freeze Period to ensure you have all data correctly. If you are trying to perform some form of Initial load and a delta load later, you will have to find a way to determine all relevant delta changes are in scope, as this is not built into the Infoporter functionality when using CSV (delta mode only works for Web services). This may require freezing certain data earlier. It’s typically possible to freeze the company org structure and position data earlier than the HR data. You therefore work with multiple freeze points, which adds complexity on the level of change management and communication but reduces the complexity of the cutover.

There is extensive benchmark data available for CSV imports that allow you to estimate how long something should take. These can be found here: https://help.sap.com/viewer/f5c753ba58814ef0ab181747824c41ed in the chapter Performance Benchmarks. This can also serve as a guidance to optimize CSV based migration for performance during cutover.

### 5.7.1 Performance Settings

As with Web Services, you will need enough background processes to export the data from SAP ERP HCM in parallel. When using CSV the SAP ERP HCM data export runtime is usually not the bottleneck. Here some considerations in EC to reduce the bottleneck:
• The import will go faster in “full-purge mode”. This implies that with the active data you also upload the history.
• In the company settings:
  o As a batch size for imports you should use 500, the default is 50.
  o As a threadpool size you should use 5, the default is 1. This will enable the parallel import of 5 CSV files at the same time.
• You should disable rules processing in Employee Central during imports. However, this requires that the fields that are filled in by these rules are filled in as part of the imports that you perform.

With these settings you can load up until 5 CSV files at the same time into Employee Central.

5.7.2 Splitting the Import File

Deciding how to split the files will depend on the quantity of data and is largely driven by the number of employees.

The limit of the file size is 50,000 records for all. If you have 30,000 employees, each with 2 addresses, you may end up with 60,000 records and the file must be split into multiple files. The same applies to EC Entities which are effective dated and where you upload records active after the Full Transmission date. Optimal performance is reached with files of 10,000 for the majority of the EC Entities. Most benchmarks are also done based upon 10,000 records.

To optimize performance, you need to evaluate the size of each file during the testing. This implies you test with a full set of production equivalent data. You need to consider the effort of splitting the files in combination with the possibility of using up to 5 jobs at the same time. Splitting can take a lot of effort during a time critical period.

• For less than 10,000 employees it’s probably easier to not split and load a single file, even if that will take more time than 5 files of 2,000 in parallel.
• Similarly, while 10,000 records provide optimal performance, splitting a file of 15,000 in 2 files may take more time than directly loading the file. This is a judgement call.
• As 20,000 records in a file upload will take more than double the time required for a file of 10,000, it is recommended to always split when you are loading more than 20,000 employees. You can then benchmark that effort and the impact on your cutover.

It will depend also on how much capacity you have available to perform the manual file splitting work. Spending sufficient time on this decision and the pros and cons will help to understand the run times and the possibilities to optimize this.

A common mistake is to assume 5,000 employees translates into 5,000 records. This is very unlikely as most data will have history or may have more than one record per EC Entity. There is for example usually more than 1 address per employee to migrate.

5.7.3 Organizational Data Migration

The quantity of Organizational objects (Type O) is typically low, and the OM structure does not change frequently. Due to this, this data could be loaded even before the Transactional Freeze Period.

5.7.4 Position Data Migration

An option to consider is whether the Position structure can be frozen earlier, together with the Org structure. While a freeze of HR Data will always be a challenge, it may be possible to start the position load before the Transactional Freeze Period.
5.7.5 Employee Data Migration

The PA Data migration typically falls within the Transactional Freeze Period. How much data needs to be migrated will depend on the number of employees and the amount of history.

Data before the “Earliest Transfer Date” is considered history and will require changes to the settings to be included. By default, data before the “Earliest Transfer Date” is not included. This is determined by the “Earliest Transfer Date” value in the Infoporter, and by default applies to all templates in the same template group.

A recommended way to optimize performance is to use 10,000 records per file and 5 files in parallel. If you are migrating more than 50,000 employees with more than 20 templates you will need multiple batches, leading to lengthy import activities. It may be required to consider working in shifts with multiple resources as this will probably take 8 – 12 hours for the loading alone, without validation, checks or dependencies.

5.8 Optimizing Initial Replication from EC to SAP ERP HCM

Replication has different requirements and works differently compared to Data Migration.

- Replication from EC Employee Central to SAP ERP HCM always uses the CPI and Web Services.
- Replication processes follows the sequence:
  - Organizational Data
  - Employee data (all templates at once)
  - Organizational Assignment Data
- Replication uses Dialog Processes and not Background Processes in SAP ERP HCM.

Once all data has been migrated to EC you perform a full initial replication from EC to SAP ERP HCM to align the SAP ERP HCM data with the EC data. After the initial replication you set up the recurring replication which will only capture changes since the last successful replication. “Successful” means that the query executed correctly, and a replication “timestamp” has been registered in the systems, even if there are errors during the replication.

Data that did not replicate correctly will stay available in the staging area until cleaned up. This works differently depending on the type of data (OM, PA and OA data).

5.8.1 Tuning SAP ERP foreground and background process types

Replication requires many foreground (ERP dialog) processes in order to run efficiently. The performance will therefore depend on the number of foreground processes available in SAP ERP HCM. When using the default settings (the SAP ERP HCM system decides how many processes to use) you will either use only half of the available foreground processes, or 10, whichever is lower. Testing will help you determine the optimal quantity. It is recommended to start with 20 dialog processes, meaning 10 of those will be used.

You can use the Operation Modes setup for Data Migration to switch the process type from Background processes to Dialog processes.

5.8.2 Replication process depending on object type

5.8.2.1 General Replication Sequence

When replicating the best sequence to follow is:

- Replicate Foundation data
  - Business Unit to Org Data
  - Division to Org Data
  - Department to Org Data
  - There may also be the need to replicate Job data
- Replicate Positions
• Replicate PA data
• Replicate OA Data (Organizational Assignments)

Depending on the type of data being replicated, the required steps are different. OM data is first downloaded to a Staging Area, and then the contents of the Staging Area are applied to the PD infotypes. It’s a two-step activity which requires specific jobs to be started individually. PA data is updated directly in the SAP ERP HCM system, but a simulation mode to test the process is available. OA data is downloaded to a buffer when executing PA replication in update mode. After PA replication is finished, you apply the OA data from the staging area to the SAP ERP HCM system.

5.8.2.2 Organizational Data and Position Data

When replicating organizational data, the data is first downloaded to a staging area, which can be reviewed before it is transferred to infotypes. Organizational Data and Positions replicate as follows:

• Report RH_SFIOM_ORG_OBJ_REPL_QUERY downloads the data from EC to buffer tables in SAP ERP HCM. You can see the content of the buffer using RH_SFIOM_VIEW_ORG_STRUC_RPRQ.

• Until this moment nothing has been changed in the infotypes. You then process the buffer contents using RH_SFIOM_PROC_ORG_STRUC_RPRQ. At this time the data in SAP ERP HCM is being updated. This also means that beyond this point, to fall back, you would have to perform a database restore, or accept the errors. It’s possible to use criteria when executing these programs for both the download and the buffer processing for example, processing department objects only.

5.8.2.3 Downloading Organizational and Position data to the buffer

Using criteria for the download (mentioned in the above subchapter) will not increase the overall speed of the process, however, the data download could already be performed for specific object types as soon as that data in EC data has been accepted as correct.

It’s safe to perform the download early, as no data is being updated in infotypes, and you can also consider applying changes to infotypes for certain objects once they have been signed-off. The risk is of course that if at a later stage there are issues with PA data you end up in a mixed state.

Note that when you download early, you will apply the data valid at the download time to the infotypes at a later stage. Changes after the download will not be applied to infotypes. These will be processed in the next execution of the download process.

If you need to restart form the download step, you can:

• Empty the buffer and restart the download (if you have the wrong objects)
• Leave the buffer as is and restart the download (if you have incorrect data for an object). It will overwrite the same entries

There are no official organization and position objects download benchmarks, however it is assumed that 50,000 objects per hour can be normally achieved. Positions will download slower than departments and business units.

5.8.2.4 Applying changes from the OM Buffer

It is possible to trigger the data transfer from the OM Buffer to PD infotypes for all object types at once, however, tests show that this process is better performed object type per object type. By doing so, it will result in less conflicts.

Conflicts can arise for example in situations such as the following: There are cases in which the update of the Business Unit leads to a minor change in all underlying positions (this depends on the configuration at the SAP ERP HCM). If all these objects are processed at the same time, you may have locking or collision conflicts (since the changes on business unit caused a change in a given position object and later on that same position is updated once again coming from the OM buffer).
Triggering the data transfer from the OM Buffer per object type means you can validate the results step by step, rather than to wait for the complete job with all objects to finish. During tests you can use the table browser with the criteria “last update date” and “user id” (that performs the update) to check what data is being affected.

If you have a lot of positions in the buffer, you could download the list, split it, save it as a variant and run multiple positions at the same time using multiple background jobs. This would only make sense with very large quantities and when you have a very short cutover period. Note that downloading and splitting the entries also takes time, and mistakes can be made. Also, the splits can cause locking errors to occur and you will need to re-run these (the typical case will be when an org object is locked due to the manager position being updated, and another position within the same org object is also being updated). It may be better to benchmark the total run, and plan accordingly, rather than to add manual steps and create the need for re-runs.

Based upon observations, updating “Departments” typically take quite some time because the changing of the OM Object leads to changes to the underlying objects. 5,000 Departments per hour can be typically achieved. Updates on Positions occur faster than on department objects. If you have 100,000 positions, assuming you have a correctly sized system, updating 20,000 positions per hour is typical. The performance per position does not increase or decrease significantly with smaller or larger volumes. Note that if the performance is a lot higher, it may be because the data to update is very limited, or errors are occurring.

5.8.2.5 Employee Data Replication

PA replication is a one-step activity. There is no download to a buffer. However, you can run replication in simulation mode. This allows you to review the replication error log without applying the changes to the database. When the PA data is applied to the database, the OA data changes are added to the OA Buffer. You run the initial PA Replication by executing ECPAO_EE_ORG_REPL_QUERY. Typically, this program is executed without filter criteria. The process will not be faster by launching multiple queries with different criteria because the query simply requests EC to start a download and CPI will split the download into batches. It may even go slower due to the criteria. Criteria in ECPAO_EE_ORG_REPL_QUERY should be used to limit the scope for testing purpose not to split the jobs.

5.8.2.6 Organizational Assignment Data Replication

Organizational Assignment Replication (OA) will create the relationship between the position and org. objects and will update Infotype 0001. In other words, it will perform the PA-PA Integration. There is no simulation mode for this. The required changes will be in a buffer, ready to apply. The buffer is filled automatically when replicating PA data. There are no tuning options for this process, other than that it runs in parallel with PA replication and is typically not the bottleneck.

After PA replication you can process the OA changes which are in the buffer and apply them to the database. Strictly speaking it is part of the cutover, but you may decide to already release the EC system for changes and usage, as all the data has been downloaded. This is a judgement call based upon the volumes and the cutover window. Either way, no new changes will be processed until you start running the recurring replication.

OA changes get processed sequentially to avoid locking conflicts. Similar to OM Replication, you could manually split them up but may have locking conflicts. This process is slower as it is a single batch process and you must wait till PA replication is done. It is not recommended to manually split this in batches to spread the load. You will need to monitor, and restart cancelled jobs which failed due to locking conflicts.

5.9 Performance Optimization for Employee Data Replication

5.9.1 Process settings

When executing PA replication, the CPI will take the total volume requested, and start downloading it to SAP ERP HCM in packages of 400 Objects. This setting is managed on the CPI and can be changed (however,
400 is a good default). If you change this, make sure this is well understood as it will change the total number of packages.

Once the first batch is downloaded, SAP ERP HCM will start applying the changes using dialog processes, while the other batches continue to download. By default, SAP ERP HCM will apply up to 10 batches in parallel, this means 10 x 400 = 4,000 employees in parallel. Therefore, if you have 16,000 employees, you will need to process 40 batches.

If the process needs to go faster and you have enough capacity in the SAP system, you need to ensure the system uses more dialog processes, by changing the settings in the bgRFC Configuration in SAP ERP HCM. This is typically done by the customer’s SAP Basis team. In the default settings SAP ERP HCM system will decide how many schedulers to use, and how many dialog processes per scheduler. This will usually be one scheduler and 10 dialog processes. You can increase this by increasing the supervisor destination (to change the server(s) used), scheduler count (on that server), or the number of processes per scheduler on that server. This depends on whether you want to distribute the load across application servers, or centralize it on one server, and whether you see a lot of messages indicating waiting for scheduler. This is only required for high volumes. Search for bgRFC Configuration on help.sap.com for a description on how to perform these changes.

A prerequisite is a correctly defined logon group and a correctly setup supervisor destination. The inbound destination can then be linked to a specific group of servers, or a logon group. You can then add more schedulers by app server or logon group if the system is under heavy load. However, if the load is light it’s better to have less schedulers otherwise they could block each other. You also need at least 3 dialog processes per scheduler on every application server in use.

The default bgRFC Gateway resource usage of “50%” means only 50% of the available free resources will be used. You can increase this if you do not need outbound resources like for the Data Replication Monitor, where also 50% is reserved by default. You can then also increase the number of Destinations per scheduler but need to ensure the number of schedulers time the number of destinations is lower than the open connections value defined in the bgRFC settings.

Note that with the default settings, adding one scheduler will add 10 processes. It is recommended to start with the default, and only start fine tuning this when the run time is too long, as you will need to test different configurations and run replication to validate that there is a real performance improvement.

Based upon observations, systems with default settings and enough dialog processes can replicate PA data for 25,000 employees per hour. This off course depends on the number of infotypes in scope and the overall SAP system performance and network performance.

5.9.2 Data Replication Monitor

If the PA Data replication still takes too much time, you can consider disabling the Data Replication Monitor. The Data Replication monitor only transfers logs for PA and OA replication from SAP ERP HCM to EC. It does not transfer information for OM replication, and therefore does not influence its performance.

All PA and OA log information can also be found in SAP transaction SLG1 which is the source of the data visible in the Data Replication Monitor. After the initial replication, you can enable the Data Replication Monitor again. While this will improve performance, you need to balance this against the effort of using SLG1 for analyzing the logs.

If you are replicating less than 20,000 employees and do not face many data errors, disabling the DRM will speed up the replication process. It may however add more effort for analyzing errors as relying only on transaction SLG1 in SAP ERP HCM takes more time than using the DRM. A recommendation is to test replication of all data with and without the DRM to measure the difference.
6 REFERENCES

SAP Help Portal
- Central location with all SAP provided manuals to use the Infoporter integration:
  - Employee Central CSV Performance Benchmarks

SAP Notes/KBA
- 2080728 – Employee Central: What is HRIS Sync?
- 2349390 – Once Daily Recurring HRIS Sync Running Multiple Times
- 2263251 – How to run a one time Full HRIS Sync - Partner

Architectural Leading Practice (ALP)
- Data Migration Strategy