Stewardship Tier: SAP RFC Extraction Overview

VERSION 1.0

DOCUMENT HISTORY

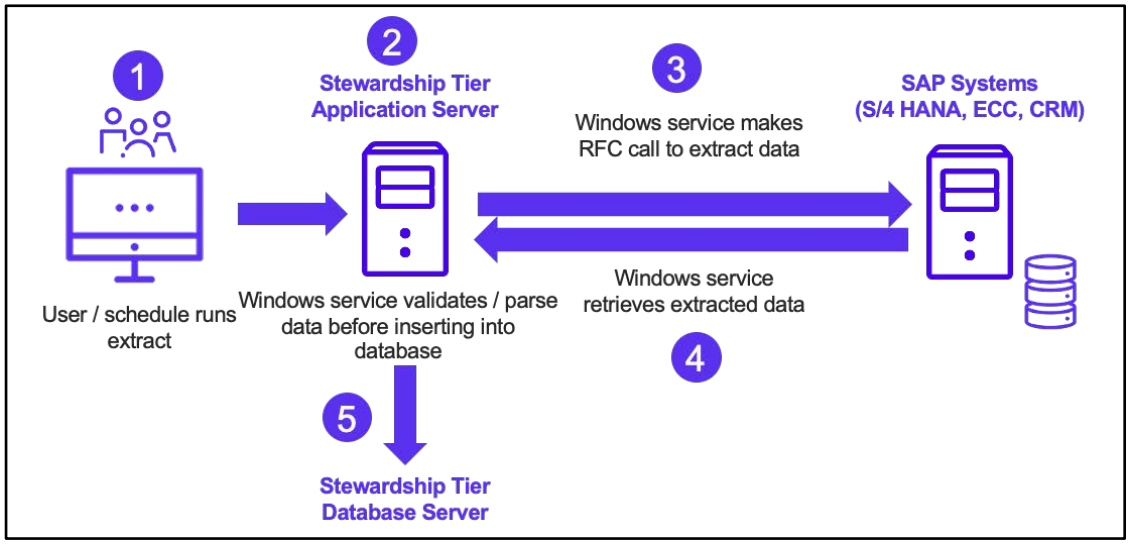
Version	Comments	Date	Changed by
1.0	1st Draft	9th November 2020	A Lund

Contents

1 Stewardship Tier: RFC Based Data Extraction Process Overview

This document provides an overview of the SAP RFC data extraction process used by the Stewardship Tier to extract data from legacy systems. The document also explains the control mechanisms available in the Stewardship Tier to manage and limit extracts.

The following diagram provides a high-level flow of the RFC extraction process:



1. User (or schedule) triggers a table extract in Collect.
2. Stewardship Tier assigns a queue and a service to the extract.
3. Based on the queue operational hours / number of threads available, the Windows service makes an RFC call to the destination SAP system to extract data.
4. The data from the RFC is retrieved based on the requested packet size, columns and specified WHERE clause.
5. The data is then validated and parsed before being inserted into a staging database.

NOTE: Steps 3 – 5 are repeated until all records are extracted.

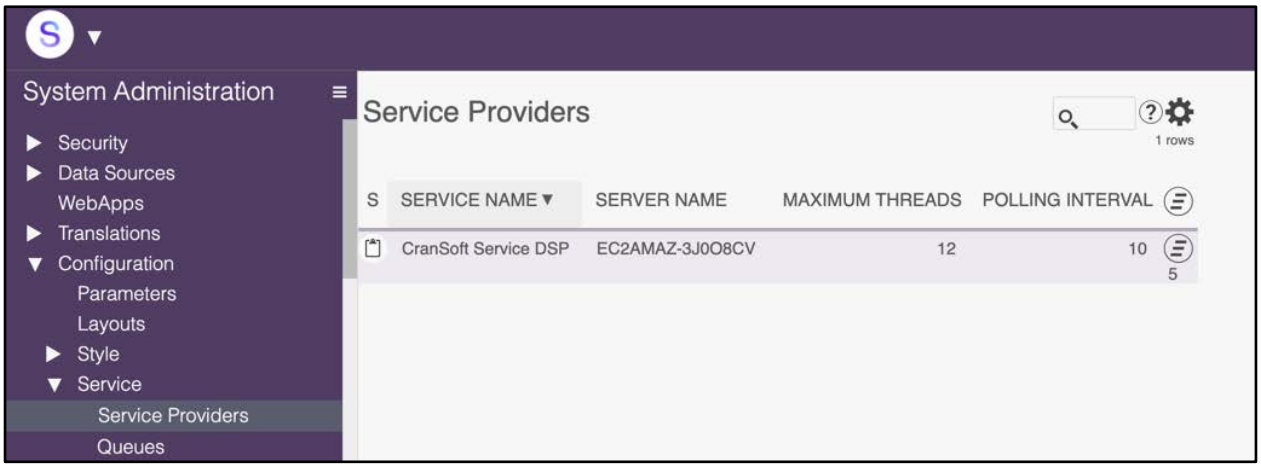
2 Stewardship Tier: Technical Overview

2.1.1 Stewardship Tier Services

Collect RFC-based extracts are executed by one or more Windows services that typically run on the Stewardship Tier application server. It is possible to distribute workloads to other servers by installing services on additional servers.

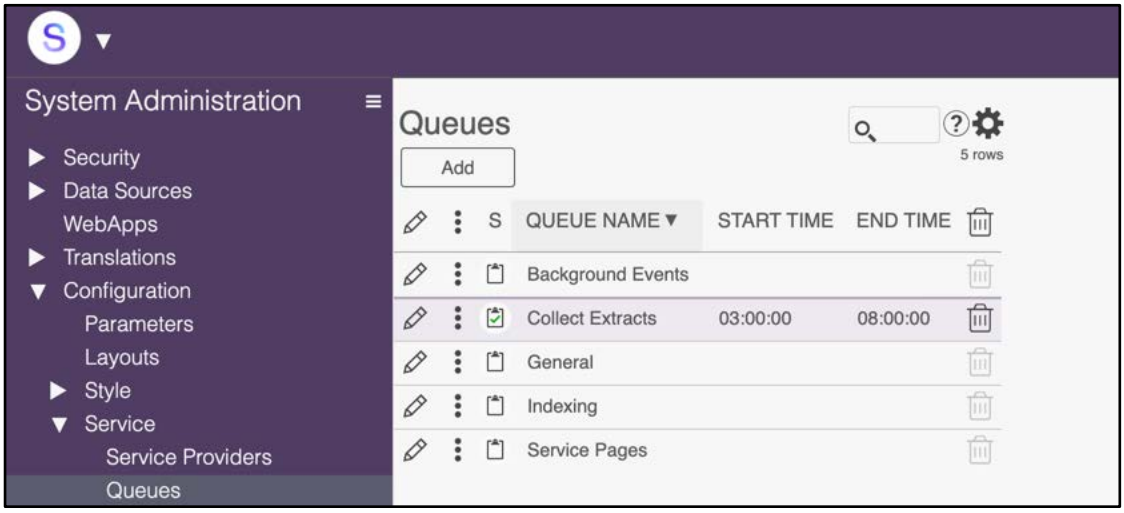
Performance Control: A service can be configured to allow for a maximum number of parallel threads (Stewardship Tier jobs). This allows resources of the server(s) on which the Stewardship Tier service(s) are running to be protected.

Security Note: Only users with Stewardship Tier Administration access and physical access to the application server can install and configure new Stewardship Tier services.



2.1.2 Stewardship Tier Queues

Queues can be created to allow better control of the processing of different job types. Queues can be assigned specific operating times.

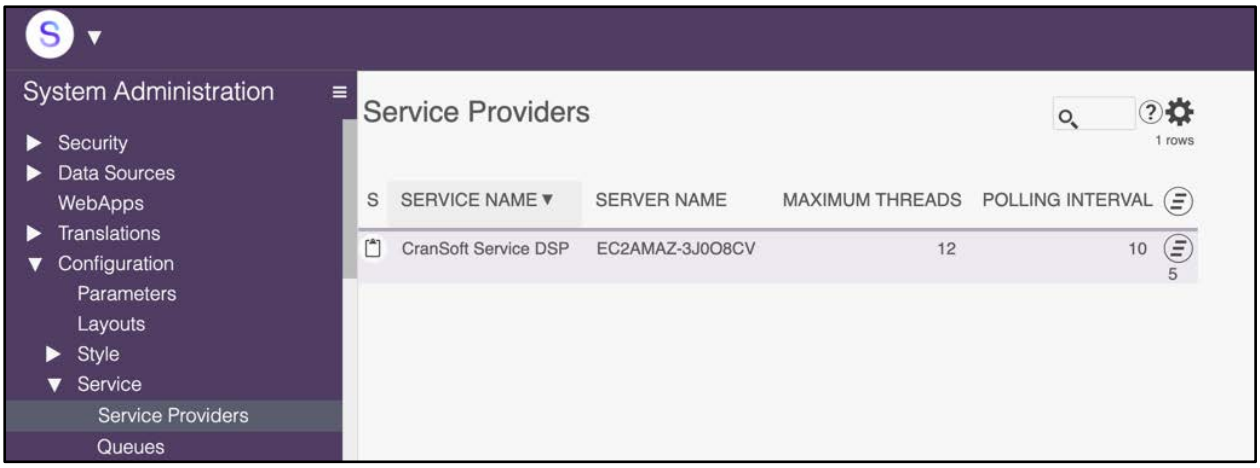


Performance Control: In the example above, jobs that are assigned to the Collect Extract queue only start running between 3am and 8am and therefore allow agreed extract processing windows to be established.

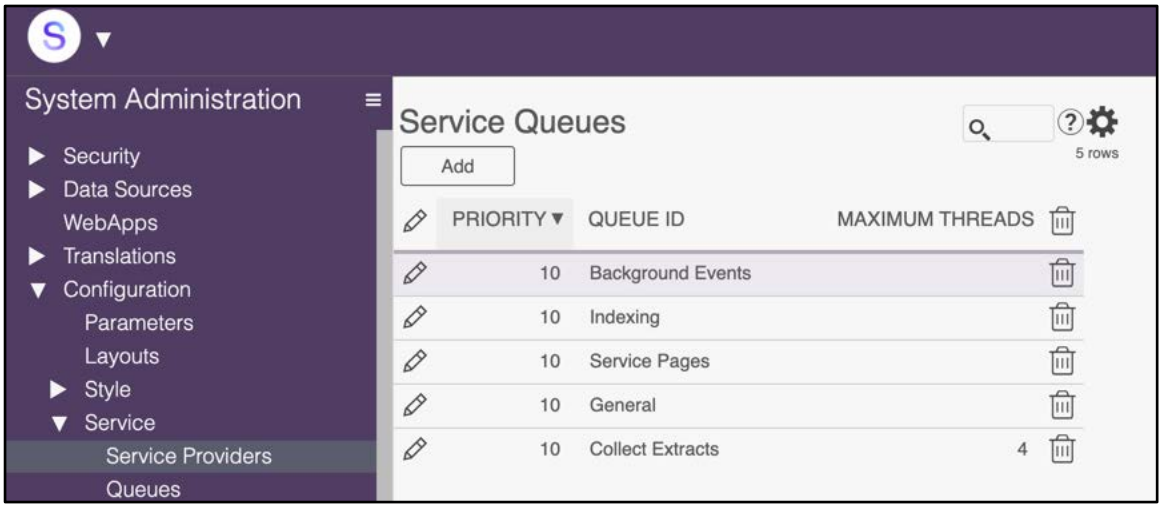
Security Note: Only users with Stewardship Tier Administration access can create/edit queues.

2.1.3 Stewardship Tier Queue Assignment to Services

One or more queues can be assigned to a service. The following example shows a service that has a maximum of 12 threads (parallel jobs) with 5 queues assigned:



The priority and maximum number of threads for a specific queue can be specified. If no maximum threads for a queue is assigned, then the queue processes the maximum number based on the service limit. If a large number of jobs are inserted into the queue, important processes may have to wait.



Performance Control: In the example above, the Collect Extract queue can run a maximum of 4 extracts at a time. It is recommended that projects work with IT technical teams to agree on the accepted number of extracts that can run in parallel and adjust this setting accordingly.

Security Note: Only users with Administration access can change the maximum number of threads.

2.1.4 Collect Extract Set-up

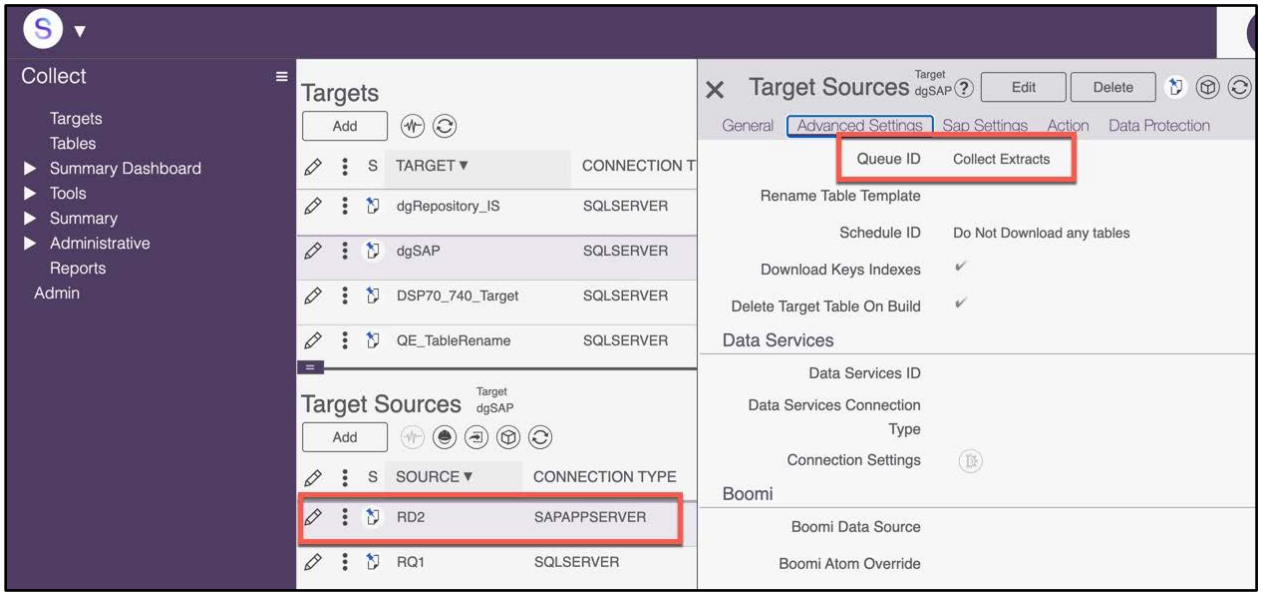
When a Collect extract is run, configuration controls which queue/service executes the job. The determination of the queue is important because it establishes the execution parameters that control when extracts are run and how many run in parallel.

The selection of which queue to use is based on a 4-level hierarchy:

1. If queue is assigned to Target Source Table -> Use that queue If no queue assigned
2. Check queue is assigned to Target Source -> Use that queue If no queue assigned

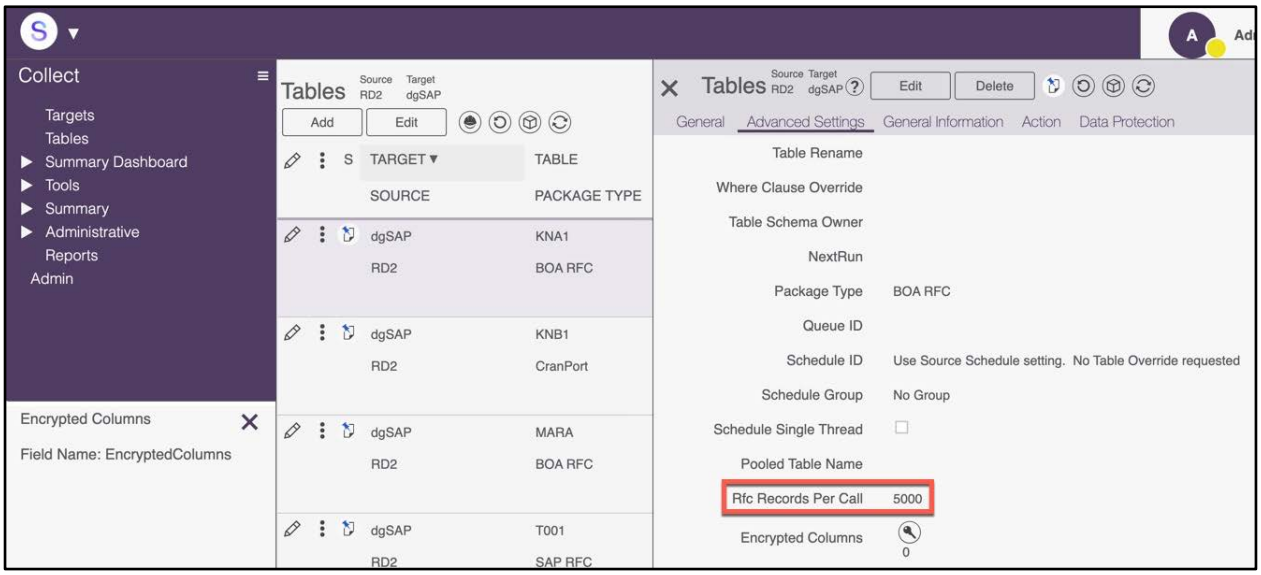
3. Check queue is assigned to Target -> Use that queue If no queue assigned
4. Use default queue defined in DSPCommon > Configuration > Modules > Collect

In the following example, the default queue has been set at Target Source level (SAP Source System). This means that all extracts from this source system use this aueue, unless an authorized user specifically changes the aueue on a given table being extracted.



Performance Control: When extracting data from a table using RFC extracts, it is possible to set a value for how many records should be extracted at a time (RFC Records per Call). This setting is important because during the extract process the RFC functions read data into the SAP application server's memory buffer before retrieving it to the Stewardship Tier. Tables that have more populated columns require more memory compared to tables with fewer populated columns and the same RFC Records per Call. There is a balance to be made regarding the number of RFC Records per Call and the time taken to perform an extract since the Stewardship Tier processes larger record packet sizes faster.

Security Note: Security groups exist that allow certain users to ONLY run extracts, whereas other users can add new tables for extract and decide whether to change the default queue for a table.



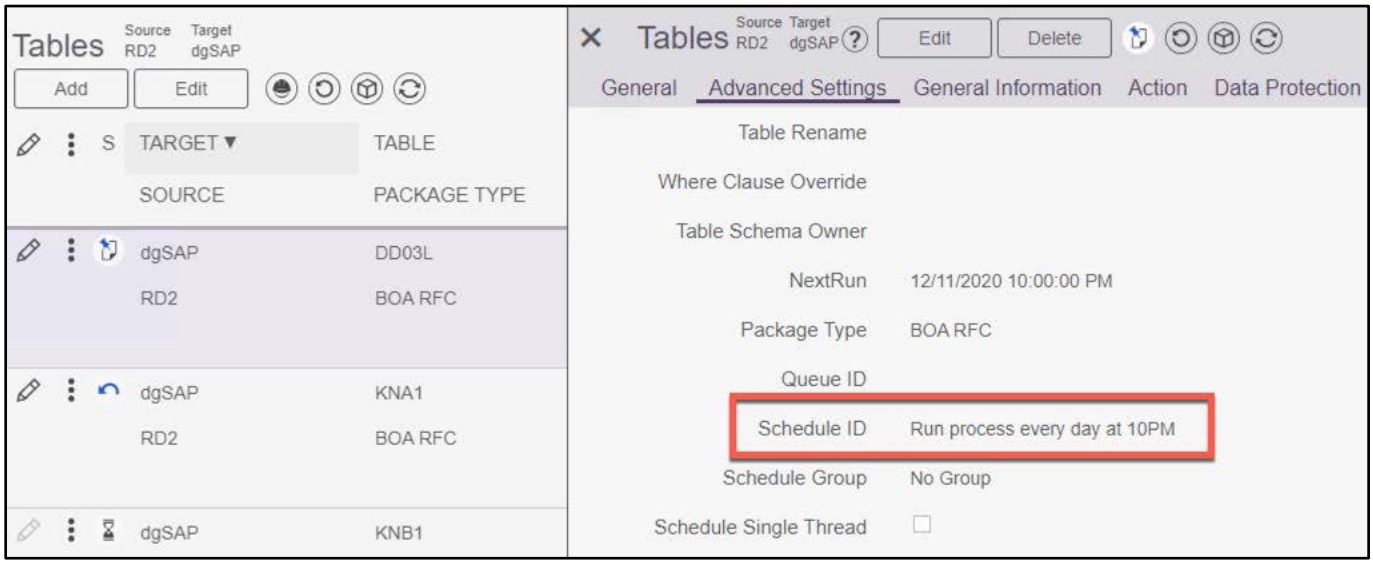
2.1.5 Extract Execution Process

Collect extracts can be triggered via the following methods:

- Schedule
- Manual Extraction of Table
- Manual Extraction of Group of Tables

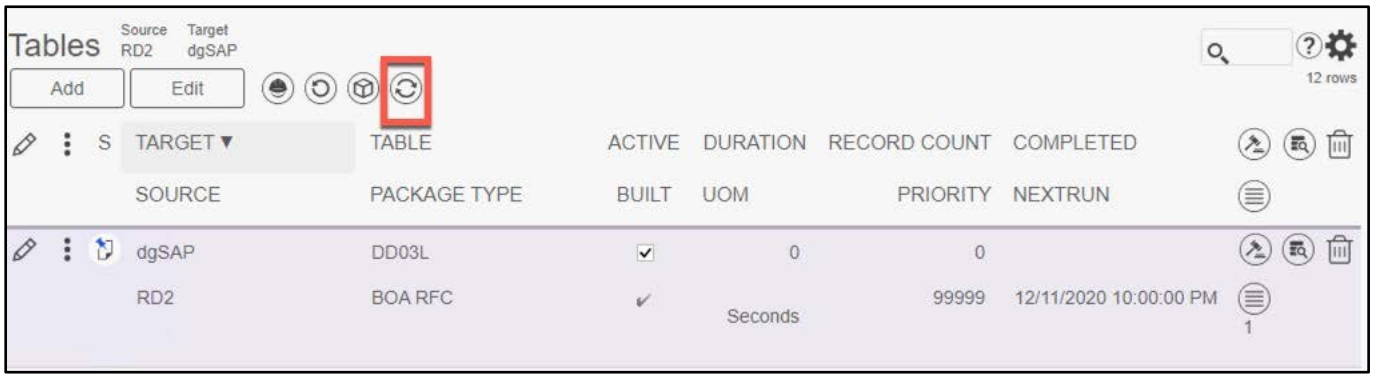
Schedule-based Extracts

Schedules can be defined in the Stewardship Tier and then assigned to individual tables. If a schedule is assigned, a Stewardship Tier service triggers the extract automatically.



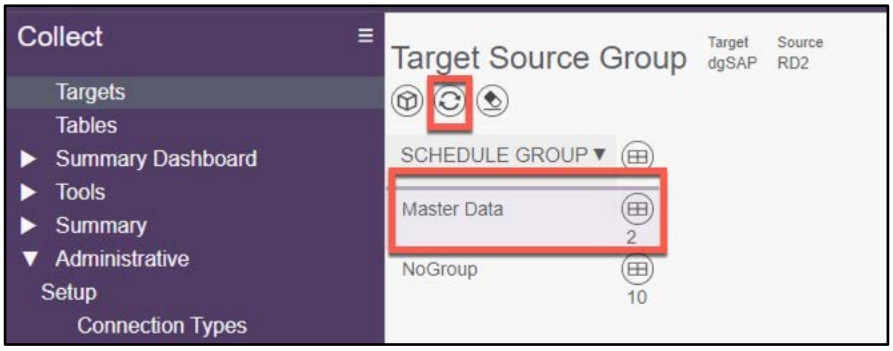
Manually Trigger Individual Table Extract

Tables can be extracted OnDemand by users with the appropriate Stewardship Tier access by selecting a table and clicking the refresh icon.



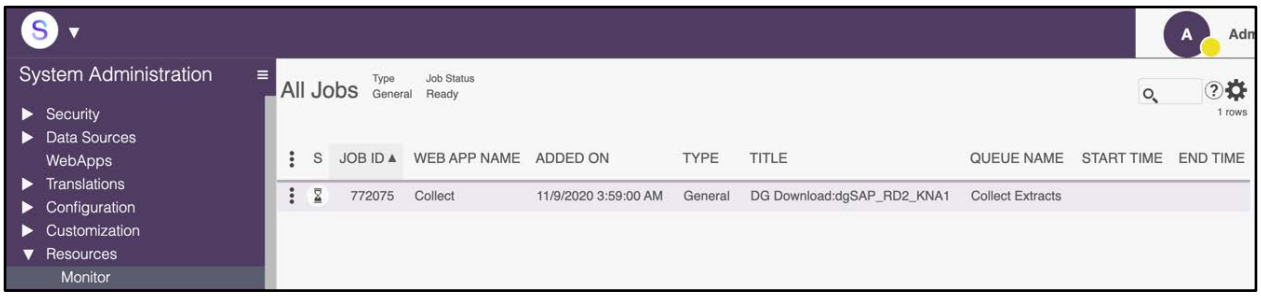
Manually Trigger Extract of a Group of Tables

The Stewardship Tier also allows tables to be grouped together via the use of schedule groups. A group of tables can be extracted OnDemand by users with the appropriate Stewardship Tier access by selecting a schedule group and clicking the refresh icon.



2.1.6 Extract Monitoring in the Stewardship Tier

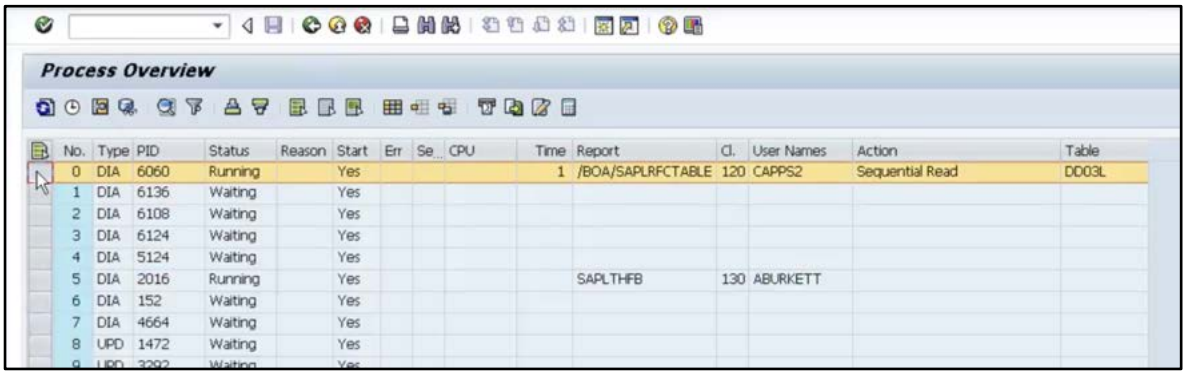
When an extract is triggered, a job queue entry is made for every table being extracted. In the following example, an extract has arrived in the job monitor and is assigned to the Collect Extracts queue; however, the hourglass symbol indicates that it is not running.



The job monitor only processes extracts if the assigned Job Queue (in this case **Collect Extracts**) is within operational hours and it only processes the maximum number of parallel extracts as defined in the queue.

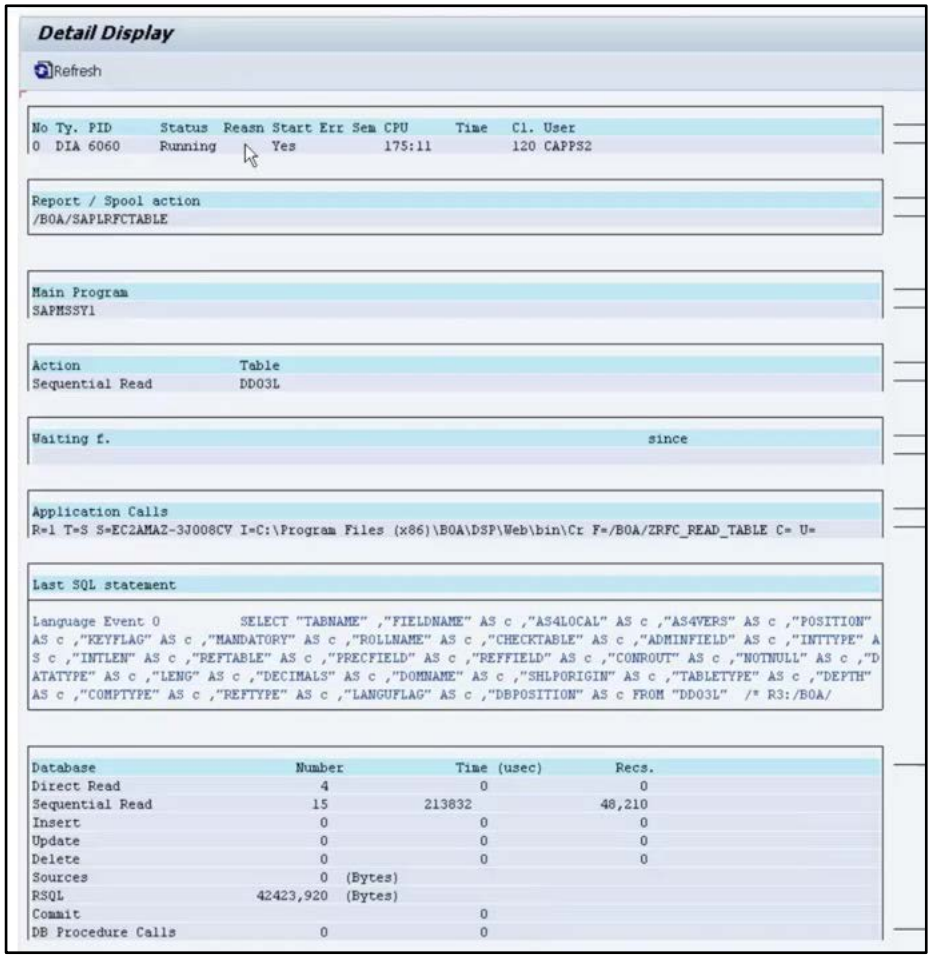
2.1.7 Extract Monitoring in SAP

The extracts performed by the Stewardship Tier call SAP dialog processes and can be monitored in SAP transaction SM50. Based on the RFC user set up for Stewardship Tier extractions, running dialog processes can be viewed.



No.	Type	PID	Status	Reason	Start	Err	Se..	CPU	Time	Report	Cl.	User Names	Action	Table
0	DIA	6060	Running		Yes				1	/BOA/SAPLRFCTABLE	120	CAPPS2	Sequential Read	DD03L
1	DIA	6136	Waiting		Yes									
2	DIA	6108	Waiting		Yes									
3	DIA	6124	Waiting		Yes									
4	DIA	5124	Waiting		Yes									
5	DIA	2016	Running		Yes					SAPLTHFB	130	ABURKETT		
6	DIA	152	Waiting		Yes									
7	DIA	4664	Waiting		Yes									
8	UPD	1472	Waiting		Yes									

Looking at the details of a running process shows the function being called, the table being extracted, SQL statement being run.



3 SAP RFC Extract Overview

The Stewardship Tier Collect module provides 3 different methods by which to extract data from an SAP system using remote function calls:

- SAP RFC
- BOA RFC
- SAP Text

There are 2 actions related to the extraction process that can be performed in the Stewardship Tier that calls functions in the source SAP System:

- Table Build – gets the table meta data
- Table Extract – extracts the data

SAP RFC Extract Method

This extract mechanism uses the SAP standard function called RFC_READ_TABLE and is limited to 512 bytes of data per record.

Build Collect Package: Calls function SAP RFC DDIF_FIELDINFO_GET to get table columns
Extract Data: Calls function RFC_READ_TABLE to extract the data

RFC_READ_TABLE Input Parameters/Tables

The following parameters and tables are passed when calling the RFC_READ_TABLE function:

- Parameter: Query – The name of the table is passed into this parameter
- Parameter: Delimiter – ' ' is supplied in this parameter to indicate no delimiter
- Parameter: No_Data – ' ' is supplied in this parameter to indicate that data should be returned
- Parameter: RowSkip – The value passed here increments by RFC Records Per Call each time a batch of records is requested
- Parameter: RowCount – The value passed here is equal to the 'RFC Records Per Call' in Collect
- Table: Options – If a WHERE Clause is supplied in Collect, it is passed into this table
- Table: Fields – The Table columns to be extracted are passed into this column

BOA RFC Extract Method

This extract mechanism uses the Syniti custom SAP function /BOA/ZRFC_READ_TABLE. This function includes some modifications to the standard SAP function RFC_READ_TABLE to support the extraction of records with up to 30,000 bytes.

- Action - Table Build: Calls function SAP RFC DDIF_FIELDINFO_GET to get table columns
- Action - Extract Data: Calls function BOA/ZRFC_READ_TABLE

BOA/ZRFC_READ_TABLE Input Parameters/Tables

The following parameters and tables are passed when calling the BOA/RFC_READ_TABLE function:

- Parameter: Query – The name of the table is passed into this parameter
- Parameter: Delimiter – ' ' is supplied in this parameter to indicate no delimiter
- Parameter: No_Data – ' ' is supplied in this parameter to indicate that data should be returned
- Parameter: RowSkip – The value passed here increments by RFC Records Per Call each time a batch of records is requested
- Parameter: RowCount – The value passed here is equal to the 'RFC Records Per Call' in Collect

- Parameter: Cross_Client - If populated with 'x' then data is extracted from all clients within the SAP system
- Table: Options – If a WHERE Clause is supplied in Collect, it is passed into this table
- Table: Fields – The Table columns to be extracted are passed into this column

SAP Text

This extract mechanism uses the Syniti custom SAP function /BOA/ZRFC_READ_TEXT. This function is used to extract SAP long texts.

- Action - Table Build: No Function call – table meta data always the same
- Action - Extract Data: Calls function BOA/ZRFC_READ_TEXT

BOA/ZRFC_READ_TEXT Input Parameters

The following parameters and tables are passed when calling the BOA/RFC_READ_TEXT function:

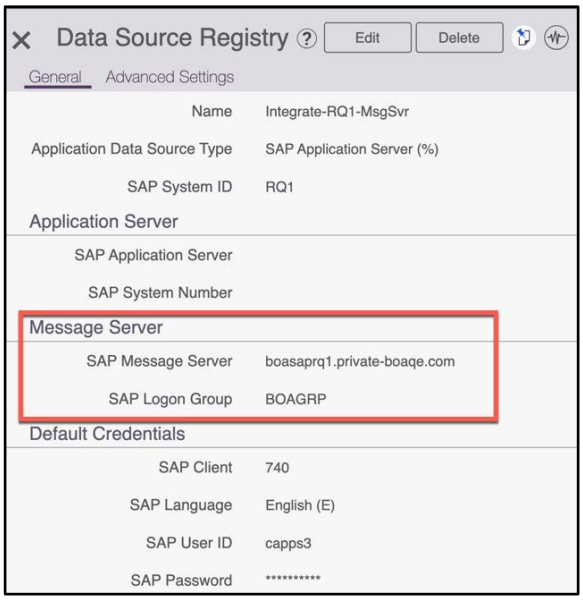
- Parameter: RowSkip – The value passed here increments by RFC Records Per Call each time a batch of records is requested
- Parameter: RowCount – The value passed here is equal to the 'RFC Records Per Call' in Collect
- Table: TEXT_LINES – based on configuration in Collect, this table is populated with the details of the Text Object, Text name, Text ID and Text Language to be extracted

4 SAP Application Performance

4.1.1 Use of SAP Message Server and Logon Groups

The Stewardship Tier also allows SAP connections to be made to an SAP Logon Group using SAP Message Server. This allows the owner of the SAP system to control the RFC execution on the SAP side.

Create separate logon groups for incoming RFC calls so that RFCs are kept separate from workprocesses of online users and thus avoid impact to dialog users. Logon groups for RFC Users can limit the maximum number of parallel jobs executed in SAP.



4.1.2 Extract Performance Thresholds

Syniti recommends that extracts are performed while the SAP application performance is being monitored. The goal of this process should be to define the agreed operating parameters for extracts that will be performed:

- Number of Parallel Extracts
- Number of Records per RFC Call
- Allowable Time Windows for Extract