Inland Revenue
Te Tari Taake

Inland Revenue

Gateway Services Build Pack: Transactional Data Service (TDS) Bulk File Feed V0.5 6 November 2017

Prepared by: Business Transformation
Date: 6 November 2017
Version: v0.5

About this Document

This document is provided to Service Providers to support the build and use of the Transactional Data Service (TDS) Bulk File Feed. It also describes the relationship with other build packs, architecture of the technical solution, schemas (file formats), and endpoints; it also provides sample file content.

This document is part of the suite of build packs that Service Providers need for implementing interfaces between their software and IR TDS.

Document Control

Document Name	Gateway Services Build Pack - Transactional Data Service (TDS) - Bulk File Feed V0.5 6 November 2017
Document Version	0.5 Draft for initial Software Developer feedback
Contact	For feedback relating to this document and/or further information please contact the Inland Revenue Software Developers Liaison Unit (SDLU) E: mailto: SoftwareDevelopersLiaisonUnit@ird.govt.nz

Contents

1 Overview

1.1 Solution overview

Inland Revenue (IR) has a suite of digital services available for consumption by our partners that support electronic business interactions with IR.

Transactional Data Services (TDS) as a business service provides the two technical services shown in the figure below:

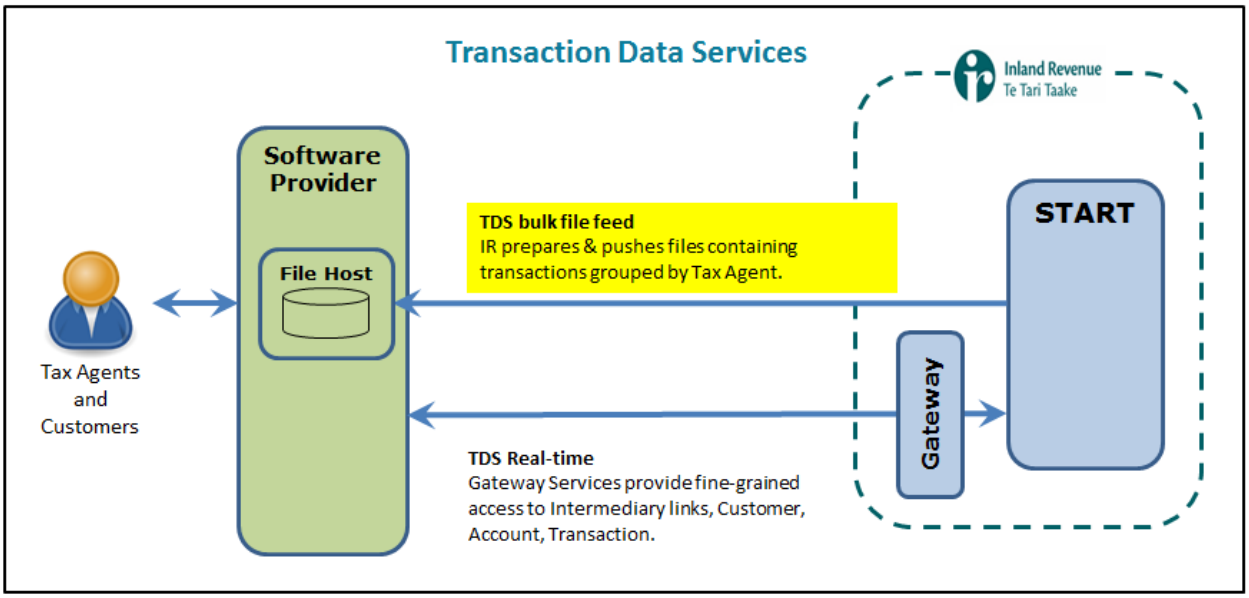

Figure 1: Transactional Data Services overview

1. This build pack focuses on the *TDS Bulk File Feed* which is an overnight file feed that pushes transaction data to Tax Agents or Customers via the Service Provider software they utilise. It is designed to cater for the high volumes of transactional data.
2. The *TDS Real Time Technical Service* is a set of web services for querying individual customers/accounts. It is intended for occasional use when the latest information is required, or information is not available from the Bulk File Feed (e.g. a new customer).

These services will only provide data for Account Types (tax types) in IR's new system, START. See Overview Build Pack for details of when particular Account type data will be available.

1.2 Intended audience

This document is intended to be used by technical teams and development staff. The reader is assumed to have a reasonable level of technical knowledge in order to comprehend the information provided. A range of technical terms and abbreviations are used throughout this document, and while most of these will be understood by the intended readers, a glossary is provided in _Intermediation use cases

The following use case is outlined here to assist Service Providers' understanding and are indicative only. The final use cases will be included in the Intermediation Build Pack.

a. Use case SUC Link service provider

SUC Link Service Provider	
User/Actors	Intermediation Services
Secondary Actor	Service Provider or Accounting Software
Description	The use case goal is to link between the Service Provider or Accounting Software and their users and to send confirmation that the link has been made.
IR systems	START
Pre-Conditions	BUC_101 Subscribe to tax data updates (refer TDS Overview BP) – includes this use case
Triggers	Request received from Service Provider or Accounting software to link a Tax Agent or Customer to an Accounting Software Provider in order to start receiving Bulk file data.
Constraints	It is expected that the Service Provider has explicit consent from the Tax Agent or Customer to create a link between them
Post-Conditions	Service Provider or Accounting Software will be sent a response from IR that a subscription link is in place between the Service Provider or the Accounting software and the Tax Agency or Customer using their software.
Use Case Scenarios	
1. Normal Flow	1. Request received by Intermediation Service 2. IR validates the OAuth token presented is for a user that has the necessary delegated authority to see all the data for the Tax Agency ID or Customer ID presented 3. Intermediation Service creates the link between Service Provider or Accounting Software and Tax Agent or Customer Account 4. IR Responds to request from Service Provider with completion status 5. Use case ends
2. Exception Flows	In all the cases below the relevant error code will be returned. Request could be rejected because of: 1. Invalid Service Provider ID or incomplete on-boarding - Please refer to the on-boarding information 2. Invalid Tax Agent or Customer ID 3. Unavailability of or internal error on the gateway services. Please contact IRD support

SUC Link Service Provider	
	4. The request message is garbled /unreadable - IR sends the user a Signal Error Response Message. 5. There is an internal error within IR IT - IR IT uses responds to user with appropriate message. 6. The User name and/or password are incorrect - IR IT responds to user with appropriate message and records the access attempt. 7. The user is valid, however doesn't have the correct permissions to use this service - For example: Is not a Tax Agent. Or, User has had their access revoked due to abuse of the service. IR records the access attempt. 8. The structure of the information in the User Request is invalid. (i.e. does not conform to XML standard) - IT records failure and responds to user with appropriate message.
3. Alternatives	For initial transition of existing service provider consumers a bulk linking process will be used. The following use cases will be detailed in the Intermediation Build Pack when complete: • Unlink Service Provider • List Service Provider Links

Appendix B - Glossary

This document is not intended for use by managerial staff or those with a purely business focus.

1.3 Related documents

All Build Packs are available on the IR BT GitHub website here:

https://github.com/InlandRevenue/Gateway-Services/wiki

The following diagram explains the relationships between the documents supporting the TDS solution:

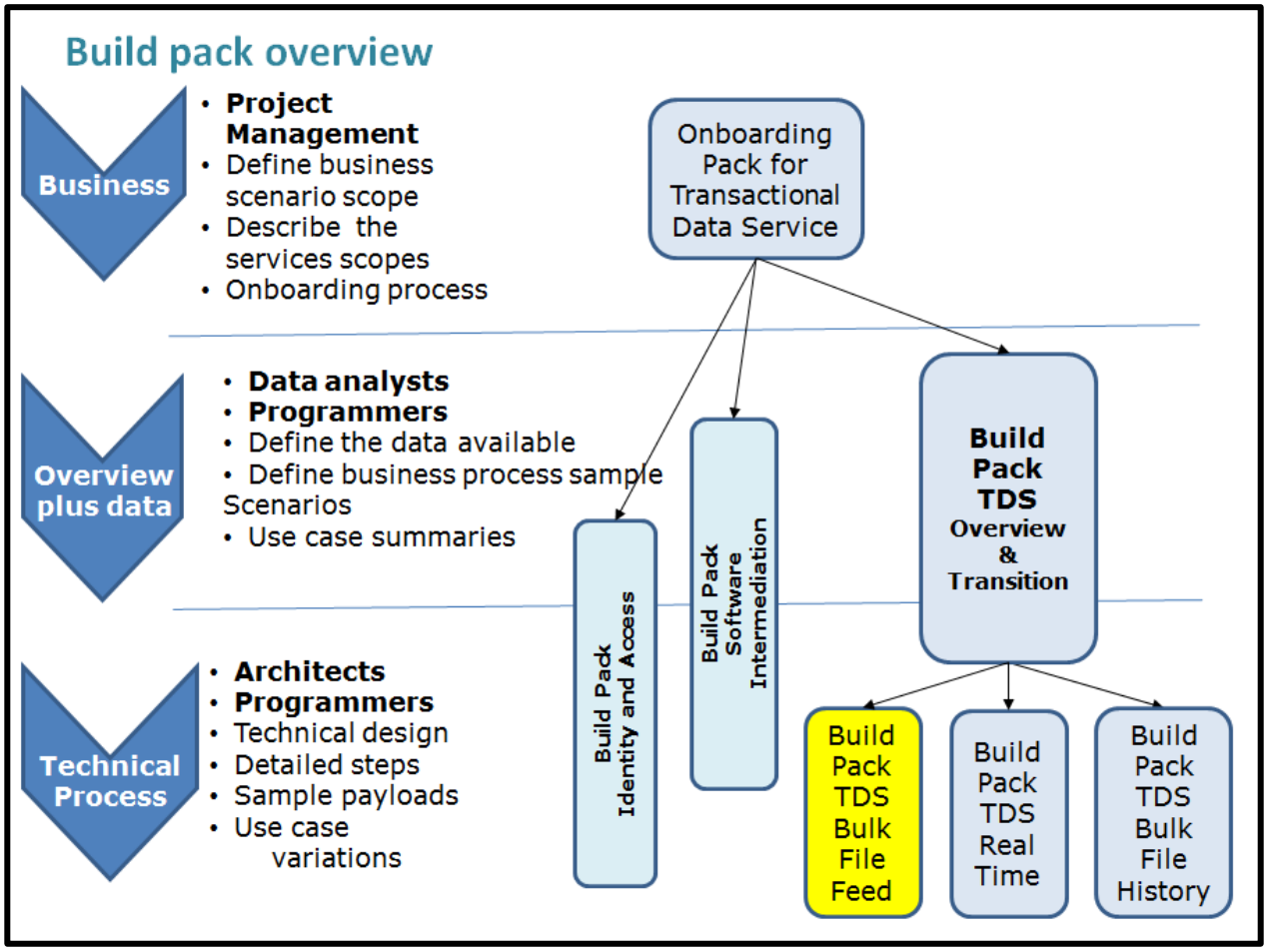

Figure 2: Onboarding and build pack structure for TDS

Name	Description

TDS – Onboarding Pack	Provides the onboarding guide for consumers of the various TDS components. Gives details of prerequisites, setup requirements, testing, contact lists, etc. It is intended to get an organisation up and running using the TDS solution. This document will not be available at the link below; instead it will be sent to Service Providers when necessary.
TDS - Overview and Transition	Describes the service components at a high level and provides an overview of the data available through TDS. Also contains information about how the component services that make up the TDS solution interact with each other.
TDS Bulk File Feed	This document
TDS Real Time Feed Build Pack	Details the technical requirements and specifications, processes and sample payloads for the TDS Real Time Feed
TDS Bulk File History Build Pack	Details the technical requirements and specifications, processes and sample payloads for the TDS Bulk File History Service
Identity and Access Build Pack	Details the Authentication mechanisms used by IR.
Software Intermediation Build Pack	Details the technical requirements and specifications for the linking of Tax Agents/Customers to Service Providers to enable these links to be used by the Bulk File Feed and Bulk file History Service.

Table 1: Related documents

2 Technical design

2.1 Overview

The TDS Bulk File Feed is intended to be used by Service Providers where large quantities of data are required. The TDS Real Time Technical Service provides similar data, but it is not suitable for high volume usage.

The TDS bulk file feed is based around a file transfer solution, where IR will send (via SFTP) information to the Service Provider on a daily (overnight) basis at the evening of each business day.

In parallel an independent second weekly cycle will also provide the same information weekly on Sunday evenings.

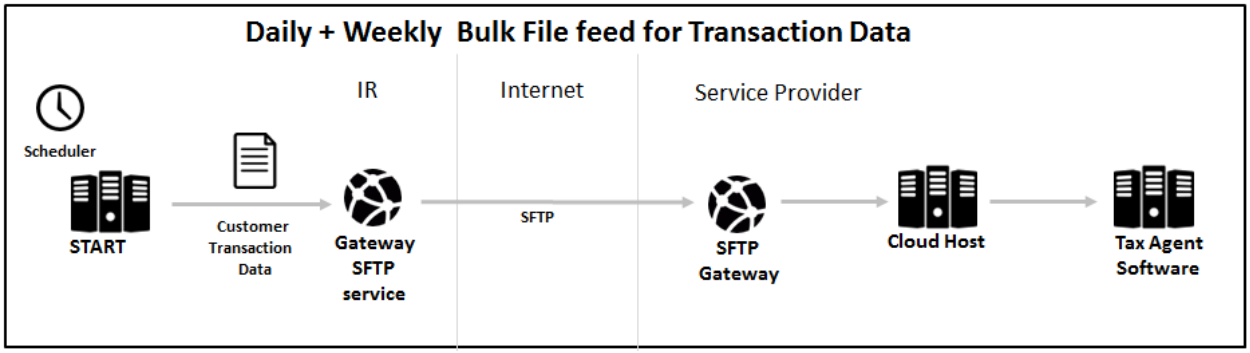

Figure 3: Daily and Weekly Bulk file feeds

Daily and weekly files are sent from START via a gateway SFTP service to the Service Provider SFTP gateway in a central cloud location from where it can be made available to their software applications and users.

Additionally Service Providers can request ad hoc manually generated files from IRD support for any period, or up to a certain date, or with all content. These files will be sent over the same SFTP channel when generated.

Each subscribing Service Provider will receive a number of zipped files containing information relating to Tax Agents that use their software product. The service also supports sending of Customer information related to large corporate customers (where there is no Tax Agent).

To determine which Customer information is sent to which Service Provider, a link needs to be established at IR between the Tax Agent or Customer and the Service Provider – this link is maintained via the Intermediation Service. This is a new web service that is being made available to support the TDS Bulk File Feed.

Please see the overview build pack for the full process and context. More details about the Intermediation Service and the Tax Agent to client linking are available in the Intermediation Build Pack.

For the purposes of this document, it is assumed that these links are in place and the business context is understood; refer to the TDS overview build pack for the broader context.

More details about the Software Intermediation Service and the Service Provider linking are available in the Software Intermediation Build Pack.

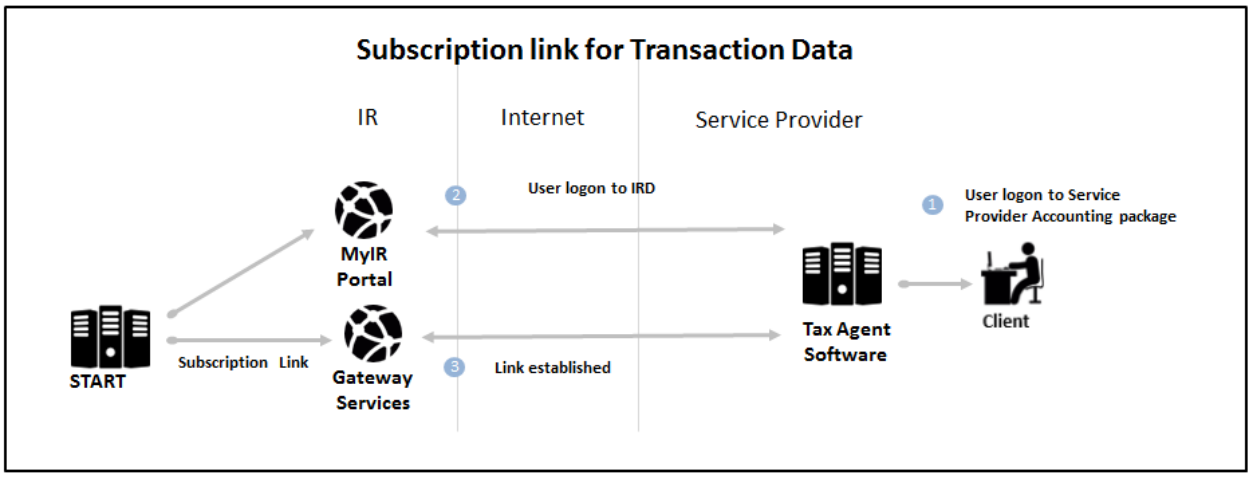

Figure 4: Subscription link for Transaction Data

The diagram above depicts the following steps:

1. The Tax Agent or other user logs into their Service Provider Software application
2. Due to accepting guidance from the package to allow bulk feed updates the user gets directed to login to MyIR portal
3. That login is then used by the software to establish a subscription link between the Service Provider and user (Tax Agent or direct Customer using their Service Provider Software application)

2.2 Transfer mechanisms

2.2.1 Connectivity for bulk file feed

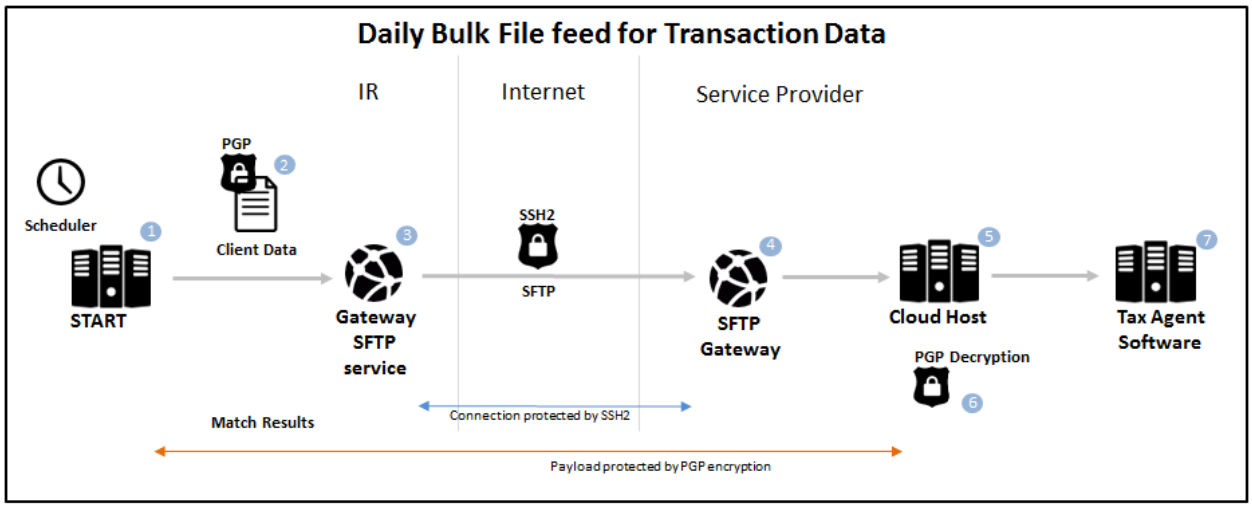

Figure 5: Daily feed - connectivity and security

The numbers above show the sequence in the path the bulk update files travel as described in the rest of this document.

Service Providers or parties planning to use the bulk file feed host an SFTP server that IRD will upload files to daily. SFTP 3.0 and SSH version 2.0 must be used. Details around this are exchanged during the onboarding phase. IRD will provide its public key from a key pair to be set up for access to the service provider SFTP site. The exact keys and their nature will be agreed during the onboarding phase. For SFTP keys IRD will prefer to be NZISM compliant and

therefore to use ECDSA keys. Where a Service Provider cannot support this RSA 2048 keys will be used. IRD expects to try to phase out non-EC keys around 2020.

PGP (as per RFC 4880) is used for payload encryption and signing—this is required due to the sensitivity of customer data being shared and especially considering the large volumes involved. IR thereby ensures that once a file is transferred to an endpoint only an authorised party can interpret it. As per PGP convention the receiver (service provider) keys are used by the sender (IRD). These PGP keys need to be 2048 bit RSA

The PGP encryption will use Advanced Encryption Standard (AES) with a 256-bit key and the PGP hashing will be done with Secure Hash Algorithm (SHA) SHA-256.

2.2.2 Connectivity for subscription

The following diagram explains the connectivity sequence protocols and security around setting up the subscription link for receiving Transaction Data:

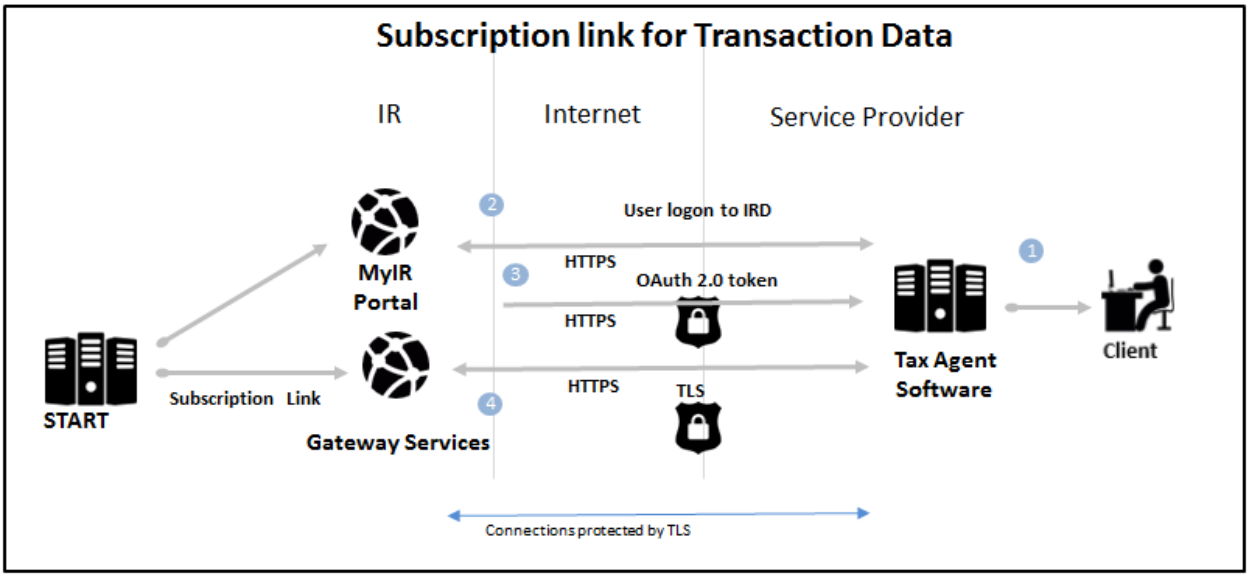

Figure 6: Subscription link for Bulk File Feed – connectivity and security

The numbers above refers to the following sequence of events:

1. Tax Agent uses their Service Provider Software application
2. The user is presented with a browser showing an IRD login the user needs to complete
3. Based on the login an OAuth token is eventually returned to the Tax Agent software
4. That token and the Service Provider's connection to IRD are used to call the Intermediation Service to link the Tax Agent to the Service Provider so that any transaction data updates for the Tax Agent are sent to that Service Provider going forward.

2.3 File structure

The files sent via SFTP are zipped files as described below.

2.3.1 Multiple ZIP files

Each daily set of information will be sent as one or more ZIP files containing customer information.

The daily increment cycle and weekly increment cycle are independent but very similar. The structural description and zip file logic and control files described below applies to both, with only filenames and timing differing. There will therefore be two types of sets of files - a set for a day and a set for a week. There will be a third set if there was a request for some manual files.

A control file will be sent containing a list of all ZIP files to be sent and the files inside each of them.

The control file serves various purposes:

- It needs to be used to know when the zip files received is equal to the intended list in the control file.
- Some Service Providers might use it to validate that all intended files in the zip files are present, other Service Providers might just rely on the PGP signing to ensure that.

The size of each ZIP file will be limited to 100 million transactions to optimise file transmission performance including retry overhead in case of failure. It is possible for Tax Agents to be spread across multiple files.

Zip files are individually PGP signed and encrypted, allowing verification of content was received and unmodified.

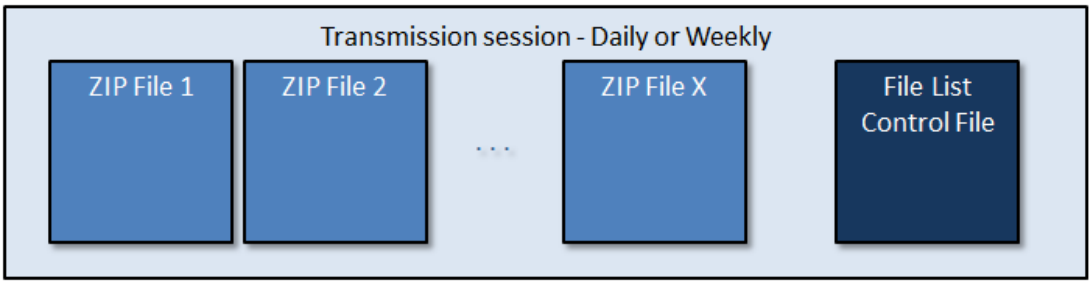

Figure 7: Zip Files plus Control file

For every occasion data is sent, whether daily as part of the daily cycle, or weekly, as part of the weekly cycle, there will be a control file listing:

- The zip file(s)
- For each zip file a list of files inside it and how many transactions and periods are included

2.3.2 ZIP file structure and control file content

Each ZIP file will contain Tax Agent files and/or Customer files. The last zip file in a set will also contain an intermediation link listing file.

Each Zip file will be PGP signed and PGP encrypted using the Service Providers Public PGP keys for signing and encryption respectively.
There will initially be one Tax Agent file with the data for all the clients linked to the given Tax Agent.

There will initially be one Customer file with the data for all the Customers linked to the given Service Provider.

To optimise performance (of encryption processing) the above files with Tax Agent information will be split across more than one file when they go beyond 100 000 transaction level elements.

All the above will initially go into one zip file, but be split across multiple zip files whenever the total size exceeds 100 million (100 000 000) transactions.

In such a scenario the data for a period of a specific Customer (or Tax Agent Client) might be split across two zip files. Along with each set of ZIP files a control file will also be present that will contain a list of zip files and the files within the ZIP file(s). This control file will also reflect the number of transactions and periods contained in each file.

All Customer data linked to a Service Provider directly, not via a Tax Agent, will be put into one Customer file for that Service Provider. That file will have a root Customers element representing the list of all Customers directly linked to the Service Provider. This file will be split into multiple files whenever it exceeds 100 000 transaction level elements.

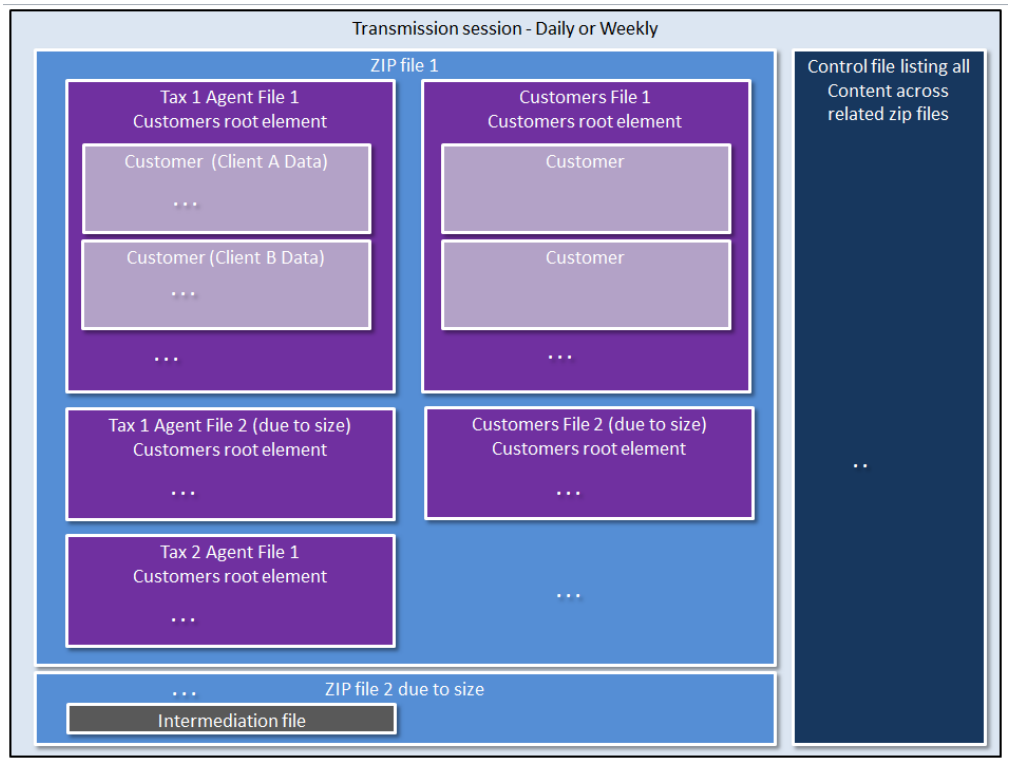

Figure 8: Zip files and their structure and summary control file

2.3.3 File structure and content and variations

Where the Service Provider linked user is not a Tax Agent or intermediary but a direct IRD Customer, there will be a single file containing all this data across all Customers. When this file becomes larger than 100,000 transaction level elements then it will be split across multiple files. Where the Service Provider linked user is a Tax Agent or other intermediary, the file(s) provided for that Tax Agent will contain data for one or more Clients of that Tax Agent/intermediary There will be one file per Tax Agent, up to the size of 100,000 transaction level elements, where after it will be split into multiple files.

A Tax Agent file will only exist in the daily set if there are changes (new transactions or new links) otherwise it will be omitted (rare in production).

Within both file types – Tax Agent or direct Customers, there is a root element called Customers and then for each Customer (Client or direct Customer) the data is structured as follows:

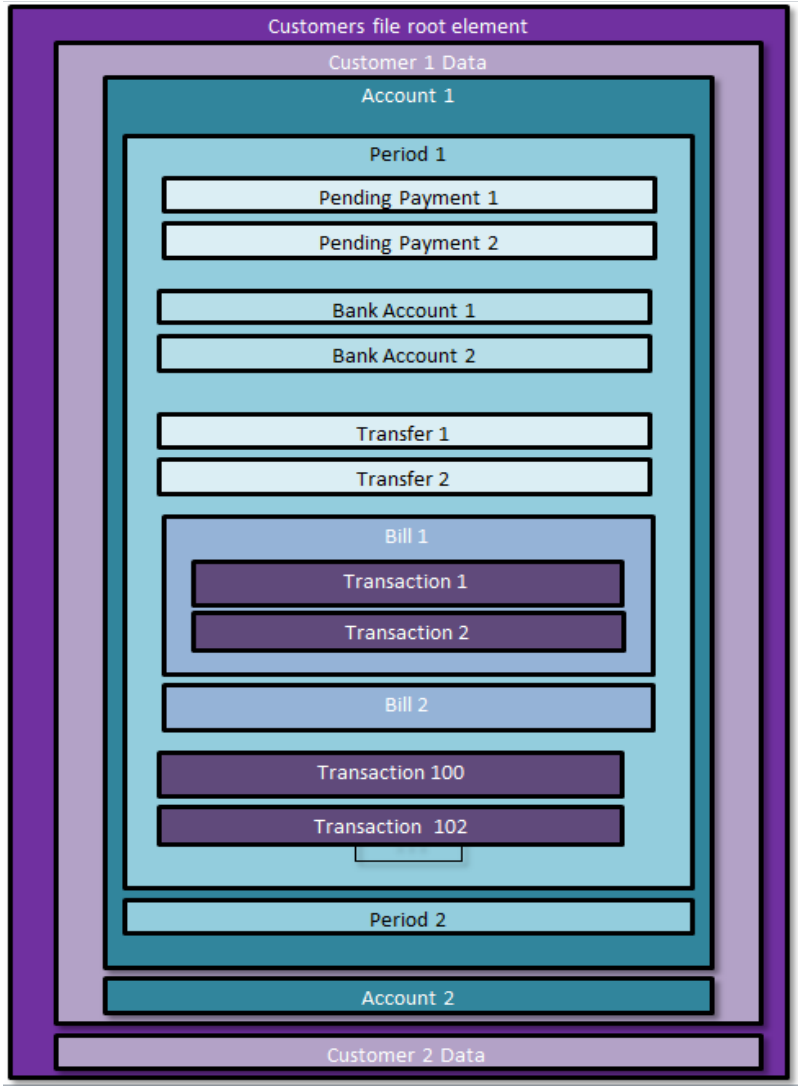

Figure 9: File content

Inside the customer section will be subsections for accounts (tax types).

Data inside accounts will be grouped by period.

For the weekly files:

a) All Customers linked to the Tax Agent will be included in the Tax Agent file. All Customers linked directly to the service provider will be included in the Customer file.
b) All Accounts linked for that Tax Agent will be included. All Accounts that exist for a directly linked Customer will be included.
c) All periods in START for each of these Accounts will be included. (See TDS Overview Build Pack for more on data available in START.) A period element will always have at least a short summary ("FilingPeriod", "Balance" and max "Activity" date attributes). This allows one to see whether the summary matches with data kept locally. (See TDS Overview Build Pack for details of the data available through TDS.)

Daily files will contain only new sections or new Transactions. Customers, Accounts, and Periods with no new Transactions or new sections will be omitted.

For the daily files:

1) Customers linked to the Tax Agent will be included in the Tax Agent file only if they contain Accounts to be included. Customers linked directly to the service provider will be included in the Customer file only if they contain Accounts to be included
2) Accounts will only be included if they contain periods to be included
3) A period element will only be included if there are new transactions to report (which might be all transactions for a new link or initial file).

In some cases the period element will contain billing groups and transactions:

a) If there were new transactions since the last file it will be included. Most transactions will be inside the respective billing group. Credit transactions will be after all the billing groups
b) If the account was recently linked to the agent, since the last bulk file feed, all transaction data for periods in START for that account will be included.
c) If the Service Provider subscribed for the data for the Tax Agent in the last day and has not received a file yet this will be a full file: All transactions for periods in START for all linked accounts and clients for the Tax Agent will be included.
d) Service Providers can request ad hoc manually generated files. This is requested via agreed support channels or Account Management. If this request has no start date the file will contain full sections.
e) If there was a technical problem at the Service Provider end, which means a full file needs to be sent, then the Service Provider may choose to unlink and relink to their Tax agent(s)/Customer(s) to invoke a new full file.

Again, the daily increment cycle and weekly increment cycle are independent but very similar. The structural description and zip file logic and control files described below almost applies to both, with only filenames and timing differing, as well as whether there are empty period summaries. There will therefore potentially be two sets of files one day a week due to cycles overlapping. There might be a third set if there was a request for some manual files.

In cases where a daily file has no transactions to report the file won't be created at all nor mentioned in the control file as omitted.

Ambiguity this might have created as to whether the agent or their clients are still linked is removed by the addition of the intermediation link listing file. The intermediation link listing file will only appear in the last zip file for a file set. It will also be listed in the control file. This intermediation link listing will provide a few types of information:

1) When did the batch run start – as a what time does this file set provide all known transactions and intermediation links?
2) What customers are directly linked to this service provider?
3) What tax agents are directly linked to this service provider?
4) For tax agents linked to this service provider, what customer accounts are they linked to?

2.4 Transfer of files – processing

On a daily basis the following is expected to happen on the site of the Service Provider SFTP endpoint/server.

1. Receive control file to help calculate processing
2. Receive one or multiple zip files from IRD
3. Validate file list in the control file matches the list of zip files received, if not, escalate to IR support.
4. Process each zip file:
    a. Decrypt zip file using agreed PGP key
    b. Verify zip file signature using agreed PGP key – if signature does not match then request IRD to resend
    c. For each file in the zip
        i. Process content of the file
        ii. Make content available to relevant data stores for users

2.5 Verifying there are no gaps between files

Max-activity timestamps on period summaries are a way to see that a period has changed since the last processed file.

Balances are similar but there are cases where there could be new transactions but no balance change, so IRD includes the max activity date in period summaries.

The primary purpose is to build trust that there are no gaps in transactions provided.

This mechanism will help highlight, together with balances changing, if an update file has been skipped – the max-activity timestamp would be different from what was stored preciously but there would be no new transactions.

The intention with the max-activity timestamp on a period is to store it and compare the stored value against the next file processed the following day.

If daily files are always processed then for any period where there is a new max-activity timestamp there will also be new transactions.

If the next file is processed and the max-activity timestamp has changed since the previously processed file, but there are no new transactions, it implies there is an update file in between that was not processed.

Max-activity timestamps on period summaries are NOT a way to validate the timestamp of the last transaction in the period. There will always be at least a small delay after the transaction before the max activity timestamp is calculated and updated, and this delay might in some cases be much larger than the typical few milliseconds.

2.6 File naming conventions

2.6.1 Name of control file listing zip files

This pattern is almost the same as for the individual zip files which is defined next below.

Format:
<frequency>_PROVIDER_<service_provider_id>_<filesequence>_<timestamp>_<environment>_CONTROL.xml

For example:
DAILY_PROVIDER_1500011034_1_201710100921548813_NZD_CONTROL.xml

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
PROVIDER	Constant	PROVIDER
<service_provider_id>	ID allocated to Service Provider by ID	Numeric 10 digit, should remain constant for a given Service Provider
<file sequence>	Next number after zip file count	1 2 3 …
<timestamp>	Time file was created yyyyMMddHHmmssffff	e.g. 201710100921548813
<environment>	IRD environment	Three letters NZD – IRD Production

Table 2 – Naming of Control File

2.6.2 ZIP files names

This pattern is almost the same as the pattern for the control files - which is defined above.

Format:
<Frequency>_PROVIDER_<service_provider_id>_<file sequence>_<timestamp>_<environment>.zip

For example:
DAILY_PROVIDER_1500011034_0_201710100921548813_NZD.zip

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
PROVIDER	Constant	PROVIDER
<service_provider_id>	ID allocated to Service Provider by ID	Numeric 10 digit, should remain constant for a given Service Provider
<file sequence>	Start with 0 for first zipfile of day and increments if there are more than one	0 1 2 …
<timestamp>	Time file was created yyyyMMddHHmmssffff	e.g. 201710100921548813
<environment>	IRD environment	Three letters NZD – IRD Production

Table 3 : Naming of Zip files

The following two file types are embedded in the zip files:

2.6.2.1 Tax agent files

Format:
<Frequency>_AGENT_<agent_id>_<file sequence>_<timestamp>_<environment>.xml

For example:
DAILY_AGENT_IRD_000000000_0_201710111532239353_NZD.xml

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
AGENT	Constant denoting this file is sent for all the Customers data linked to a Tax Agent or other Intermediary	AGENT
<tax_agent_id>	ID allocated to Tax Agent by IRD	Numeric 10 digit
<file sequence>	Start with 0 for first zipfile of day and increments if there are more than one	0 1 2 …
<timestamp>	Time file was created yyyyMMddHHmmssffff	e.g. 201710100921548813
<environment>	IRD environment	Three letters NZD – IRD Production

Table 4 : Naming of Tax Agent files

2.6.2.2 Customer file

Format:
<Frequency>_CUSTOMER_<file sequence>_<timestamp>_<environment>.xml

For example:
DAILY_AGENT_IRD_000000000_0_201710111532239353_NZD.xml

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
CUSTOMER	Constant denoting this file is sent for all the Customers data linked to a Tax Agent or other Intermediary	CUSTOMER
<file sequence>	Start with 0 for first zipfile of day and increments if there are more than one	0 1 2 …
<timestamp>	Time file was created yyyyMMddHHmmssffff	e.g. 201710100921548813
<environment>	IRD environment	Three letters NZD – IRD Production

Table 5 : Naming of Customer Files

2.6.2.3 Intermediation file

This pattern is almost the same as the pattern for the control and zip files - which is defined above.

Format:
<Frequency>_PROVIDER_<service_provider_id>_<file sequence>_<timestamp>_<environment>_INTERMEDIATION.xml

For example:
DAILY_PROVIDER_1500011034_0_201710100921548813_NZD_INTERMEDIATION.xml

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
PROVIDER	Constant	PROVIDER
<service_provider_id>	ID allocated to Service Provider by ID	Numeric 10 digit, should remain constant for a given Service Provider
<file sequence>	Start with 0 for first zipfile of day and increments if there are more than one	0 1 2 …
<timestamp>	Time file was created yyyyMMddHHmmssffff	e.g. 201710100921548813
<environment>	IRD environment	Three letters NZD – IRD Production
INTERMEDIATION	Constant	INTERMEDIATION

Table 6 : Naming of Intermediation Files

2.7 Sample payloads

This service has no incoming requests, please refer to the intermediation and authentication related build packs.

The picture below is an extract of what files can contain; please see the attached sample files below for the full context.

Note: The samples here are indicative only and still subject to change.

```
<?xml version="1.0"?>
- <Customers xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  - <Customer IRD="123088514">
    - <Account Conversion="2018-04-01T00:00:00" MaxActivity="2017-10-31T16:06:21.08"
      Balance="211020.91" Cease="9999-12-31" Commence="2017-01-01" ID="123-088-
      514-IPS005" FilingFrequency="IPSMO" Type="IPS">
        - <Period Balance="1.25" FilingFrequency="IPSMO" Activity="2017-10-
          31T16:06:21.08" Credit="-363.55" Other="0.00" Interest="6.74" Penalty="25.06"
          Tax="333.00" End="2017-05-31" Begin="2017-05-01" FilingPeriod="2017-05-31">
            - <Bill Balance="1.25" Credit="-363.55" Due="2017-06-20" BillID="1218516992">
                <Tx Type="PNLMNL" LinkID="0" Effect="2017-10-31" Post="2017-10-
                  31T16:06:20.6" TransID="2019706880" Amt="1.25"/>
              </Bill>
          </Period>
      </Account>
    </Customer>
</Customers>
```

Figure 10 : Sample file

2.8 Schema

File Type	Reference
Basic Simple types	Common.v1.xsd
Structure	TransactionBulkList_ v0.1.xsd

Table 7 : Schema

2.9Samples

Note: The samples here are indicative only and still subject to change.

File type	Reference
Zip and control files	TDS Examples 2017-11-02.zip

Table 8 : Samples

3 Onboarding

To onboard a Service Provider to consume the Bulk File Feed service, the following information is required. For further information see TDS Onboarding Build Pack.

3.1 Service provider information required

- Destination SFTP server details including:
  - SFTP server domain name (if available)
  - SFTP server IP Addresses (these addresses will be whitelisted by IR)
  - Destination directory details
- SSH keys for connecting to the SFTP server.
- PGP public certificate to enable payload encryption.

3.2 IR information

- All IR's SFTP connections will originate from the following IP addresses:
  - Test Systems 222.153.203.33
  - Production 222.153.202.39
  - Production(DR) 222.153.202.33
- Test data information
  - See examples above and future build packs

4 Use cases and process

4.1 Process model

See the overview build pack for a more detailed explanation of the stages of TDS usage, summarised below.

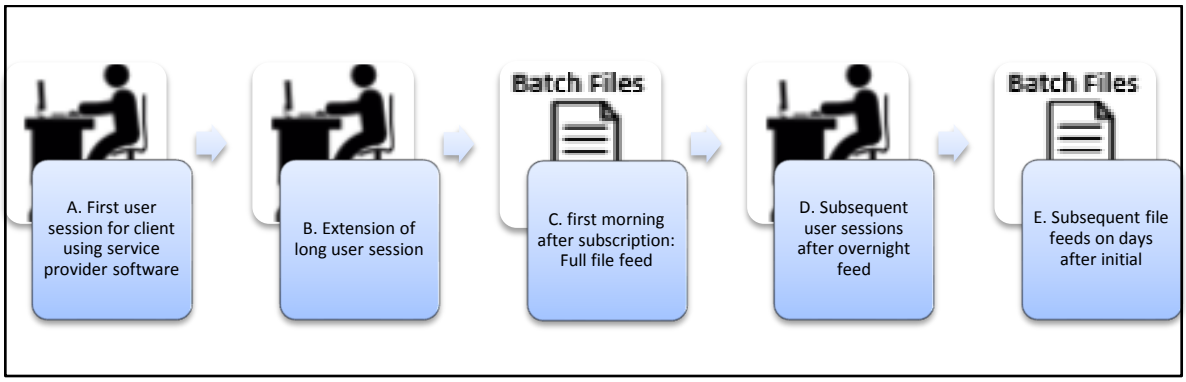

Figure 11 : Overall process

Below is the overall process model for using the Bulk File Feed:

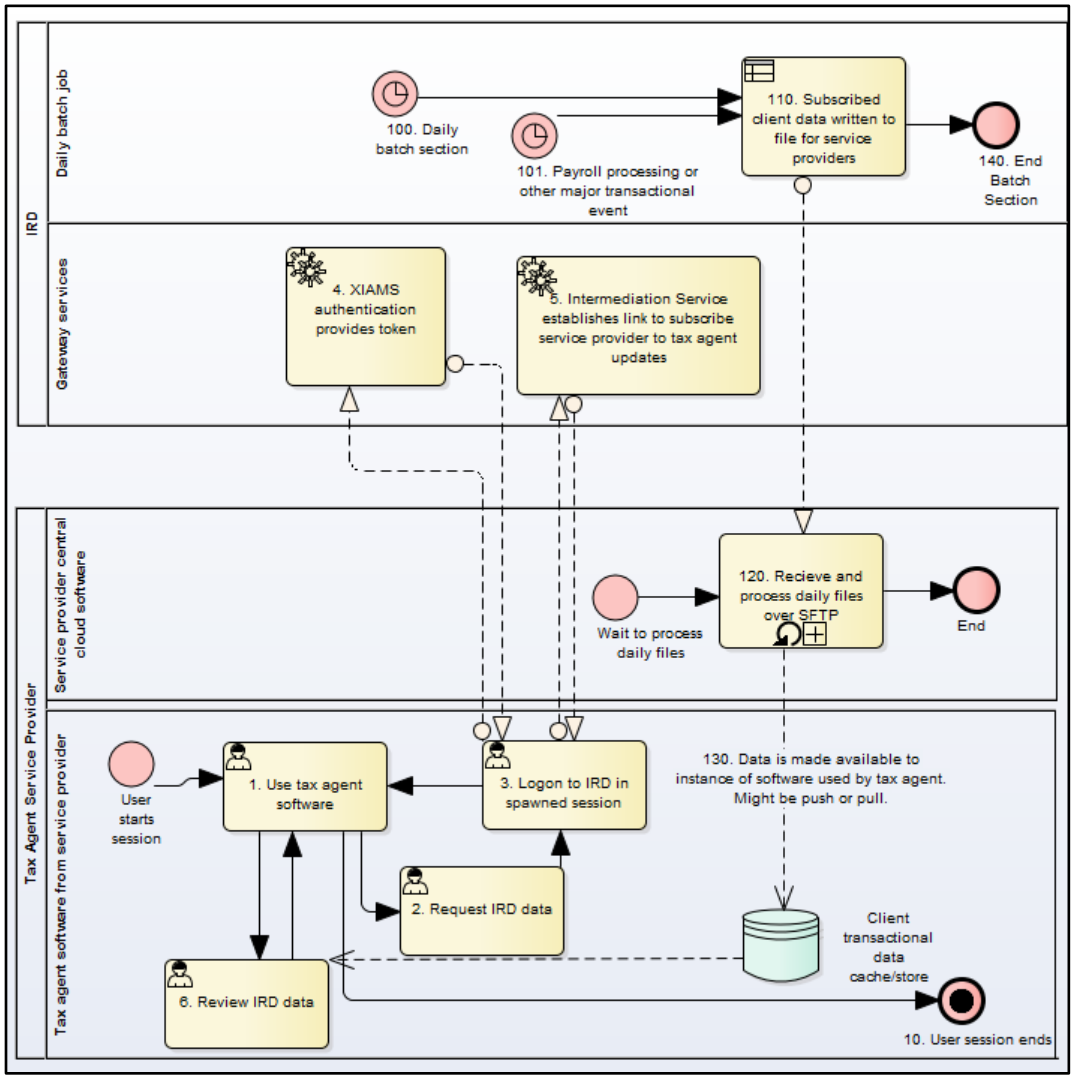

Figure 12: Process model for TDS Bulk File Feed

A. Service provider and IR setup SFTP path and exchange credentials during onboarding

1. Prerequisite (not shown on process model). See the TDS Onboarding pack for more information

B. User session for agent or customer using service provider software

1. User signs onto Tax Agent software from service provider
2. User decides to request access to some IR tax data
3. Service Provider software lets user starts an IR session by logging onto the IR site
4. As part of this logon sequence an OAuth token is returned for use in further calls to the Gateway Services. See the Identity and Access Build Pack for more information
5. The Service Provider software then uses this token in a call to the Intermediation Gateway service to request a link between the Service Provider software and the user (whether Tax Agent or direct Customer). This link subscribes the Service Provider to getting updates on transaction data for this user.
10. User might do other work in the Tax Agent software and eventually logs off and terminates session

C. First morning after subscription: Full file feed

100. During the IR overnight processing runs where most updates are done, towards the end of the sequence a job kicks off to update subscribed Service Provider software with the bulk feed.
101. As a variation of this process, it might be kicked off after a large processing event like payment runs. Outside such events very little information change occurs to justify a file feed or real time call.
110. For each Tax Agent or Customer that is subscribed to by Service Providers their data gets written to file and aggregated to send to the Service Providers.
120. The Service Provider receives the files on the SFTP location registered for that purpose. The SFTP session uses exchanged keys to establish the underlying SSH connection. The Service Provider might wait for the last summary file to be sent to trigger processing of the content. The zipped content is unzipped and the individual files are decoded using the exchanged PGP keys.
130. The content of each file is made available to the respective client. This might involve any combination of central storage, remote push to remote servers or desktops, being made available for query by remote server or desktop software
140. The job section in the daily batch run terminates.

D. Subsequent user sessions after overnight feed

1. User signs onto Tax Agent software from Service Provider
2. User can access tax data from store updated through overnight feed
3. If user needs to drill down into some data or needs to access changes to data since evening then real time Gateway Services can be called – see separate build pack for TDS Real Time Feed.
10. User might do other work in the Tax Agent software and eventually logs off and terminates session

E. Subsequent file feeds on days after initial

4.2 Use case overview

As described in the Overview Build Pack the transaction data services provide data into an overall business process that is controlled by the Service Provider software and its user. The use cases therein are diverse (and therefore neither exhaustive nor meant to represent every combination of process that the Customer or Service Provider may execute) and described in broad terms in the overview build pack– the business/organisational use cases denoted below as yellow or blue.

The corresponding technical steps inside them to integrate with IRD are defined as systems use cases with corresponding numbers SUC<nnn> below. The grey ones are covered in other Build Packs as identified in the Table below. Below this diagram is a table explaining this further.

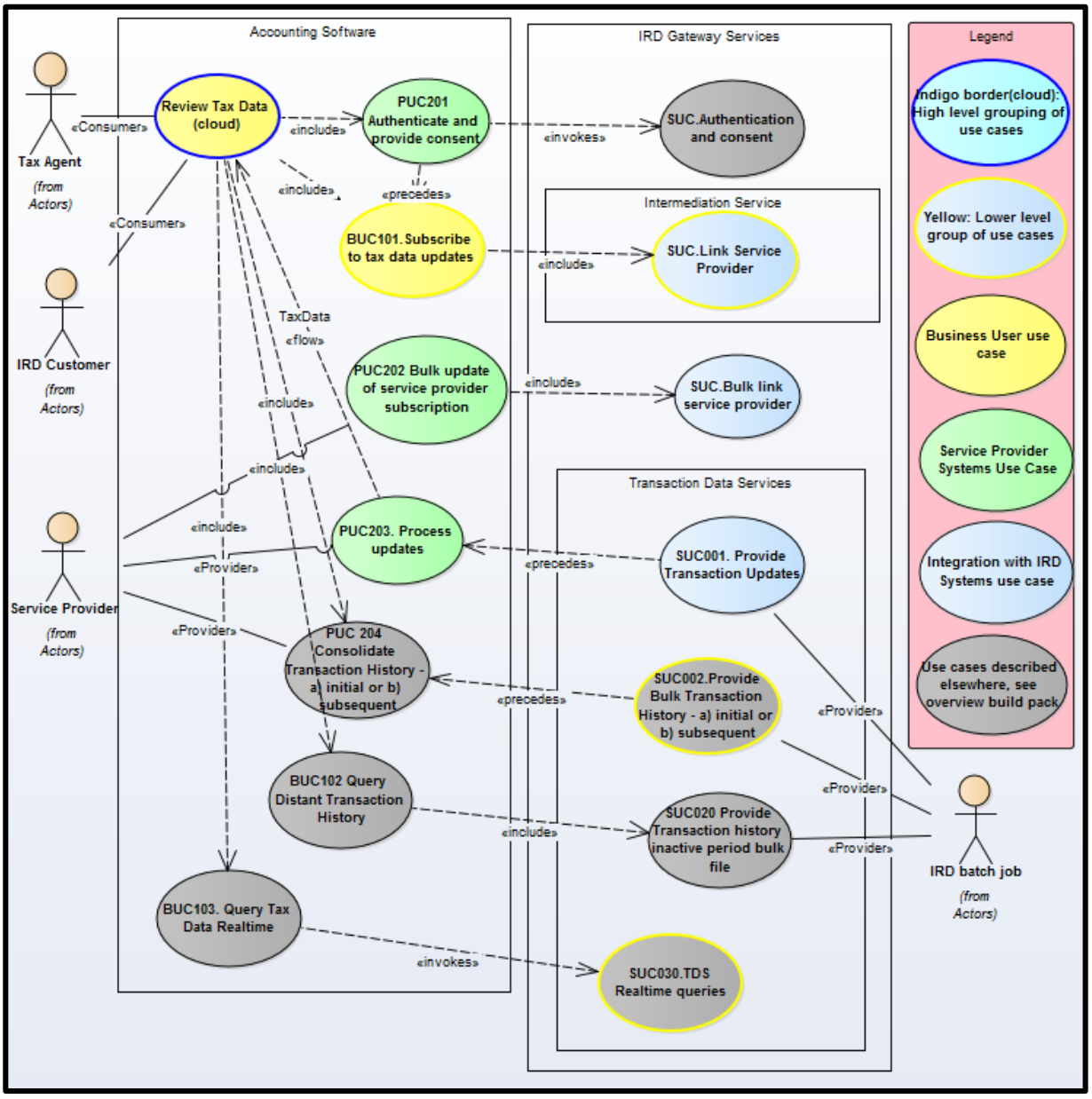

Figure 13: Use case overview

Use cases have been classified into the following types:
(Note: High level use cases are broken down in the Build Pack concerned)

Use case Group	Description	Colour	Use Case	Build Pack
Tax Agent /Customer	Use cases specific to the Customer / Agent point of view	Yellow	BUC101 Subscribe to tax data updates	TDS Overview BP
			BUC102 Consolidate Transaction History	TDS Overview BP
			BUC103 Query Tax Data real-time	TDS Overview BP
Service Provider	Some systems use cases on the Service Provider side are not user driven and broader than the integration with IR	Green	PUC201 Authentication and consent	TDS Overview BP
			PUC202 Bulk update of service provider subscription	TDS Overview BP
			PUC203 Process updates	TDS Overview BP
			PUC204 Consolidate Transaction History - a) initial or b) subsequent	TDS Overview BP
Systems use cases	The corresponding technical steps inside the above use cases which are required to integrate with IR	Blue	SUC Authentication and Consent	Identity and Access BP
			SUC Link Service Provider	Software Intermediation BP
			SUC Bulk Link Service Provider to seed the IRD START system	TDS Overview BP, see PUC202
			SUC001 Provide Transaction Updates	This document
			SUC002 Provide Bulk Transaction History - a) initial or b) subsequent	TDS Bulk File History Build Pack
			SUC020 Provide transaction history inactive period bulk file	TDS Bulk File History Build Pack
			SUC030 TDS real-time queries	TDS Real-time Build Pack

Table 9: Use cases and their relevant documentation

In the TDS Overview Build Pack sample process scenarios provide some organisational/business summary goal context.

4.3 Use Case SUC001 - Provide Transaction Updates

Use Case SUC001 Provide Transaction Updates	
User/Actors	IR Transaction Data Services
Secondary Actor	Service Provider or Accounting Software
Description	The use case goal is to retrieve data for all Tax Agents/Customers linked to the Service Provider or Accounting Software, format and package it as described in chapters above and send via SFTP to the Service Provider or Accounting Software central location
IR systems	START
Pre-Conditions	Service Provider or Accounting Software is onboarded and therefore authorised to communicate with TDS (See the TDS Onboarding Build Pack) A link / subscription has been established between the Service Provider and Tax Agent and / or Customer Account identifying which entities' tax data to send to the Service Provider and how to group it. (See the TDS Overview Build Pack)
Triggers	IR overnight batch processing runs for all business days at their evening. It completes financial updates and then 1. executes the daily update cycle 2. if there are manual updates and they are scheduled then those are also executed IR weekly batch processing runs on Sunday evenings. It 3. executes the weekly update cycle Each of the three above effectively repeats this use case
Constraints	It is assumed that the Service Provider has explicit consent from the Tax Agent or Customer to receive their data through the Bulk File Feed
Post-Conditions	The updates have been received by the Service Provider or Accounting Software and PUC203 can be completed by the Service Provider. (See TDS Overview Build Pack)
Use Case Scenarios	
1. Normal Flow	IR daily batch processing for Service Provider subscription generates PGP signed and encrypted ZIP files. These files are sent to the Service Provider at a Service Provider SFTP location. This subscription link will cause a file with an initial full data set and then daily updates to be added to any zip files the Service Provider is receiving from IR daily. 1. For each of Daily / Weekly / Manual cycle as applicable: 2. Iterate through on-boarded Service Providers and their linked Customers as well as their linked Tax Agents and the clients linked to those Tax Agents: 3. Data is formatted as described above in this document and in the TDS overview build pack 4. Data packaged in zip files as described above in this document 5. File(s) sent to Service Provider or Accounting Software central location via SFTP as described above in this

Use Case SUC001 Provide Transaction Updates	
	document 6. This system use case ends Service Provider use case PUC203 then completes at each Service Provider The steps after the SFTP transfer completes depend on the Service Provider or Accounting Software. Once all the file content has been transferred the Service Provider will process the file content and makes it available to the software instances used by the users of the requesting organisation
2. Exception Flows	If IRD plans a maintenance outage and will not be able to do a daily batch run Service Providers will be notified through agreed communication channels established during onboarding If IRD is unable to connect to the Service Provider SFTP site the system will escalate for the IRD support team to contact the Service Provider and agree a plan to restore connectivity and reschedule transmission If SFTP transmission fails it escalate to the IRD support team to contact the Service Provider and agree a plan to troubleshoot and restore reliable connectivity and reschedule transmission If all files are sent but during PUC203 processing thereof the Service Provider finds there are issues like file corruption they will contact IRD support over channels agreed during the onboarding process to agree a plan for troubleshooting and rescheduling transmission or requesting new manual files.
3. Alternatives	When a bulk feed to a Service Provider is delayed there is a limited ability to compensate by users needing updates urgently requesting it via the real time feed. This needs to be negotiated and monitored by both the Service Provider and IRD support to not generate volumes that degrades the overall experience for Service Provider users or those connecting to IRD.

Table 10 : Use Case SUC001

5 Appendix A— Intermediation use cases

The following use case is outlined here to assist Service Providers' understanding and are indicative only. The final use cases will be included in the Intermediation Build Pack.

b. Use case SUC Link service provider

SUC Link Service Provider	
User/Actors	Intermediation Services
Secondary Actor	Service Provider or Accounting Software
Description	The use case goal is to link between the Service Provider or Accounting Software and their users and to send confirmation that the link has been made.
IR systems	START
Pre-Conditions	BUC_101 Subscribe to tax data updates (refer TDS Overview BP) – includes this use case
Triggers	Request received from Service Provider or Accounting software to link a Tax Agent or Customer to an Accounting Software Provider in order to start receiving Bulk file data.
Constraints	It is expected that the Service Provider has explicit consent from the Tax Agent or Customer to create a link between them
Post-Conditions	Service Provider or Accounting Software will be sent a response from IR that a subscription link is in place between the Service Provider or the Accounting software and the Tax Agency or Customer using their software.
Use Case Scenarios	
1. Normal Flow	6. Request received by Intermediation Service 7. IR validates the OAuth token presented is for a user that has the necessary delegated authority to see all the data for the Tax Agency ID or Customer ID presented 8. Intermediation Service creates the link between Service Provider or Accounting Software and Tax Agent or Customer Account 9. IR Responds to request from Service Provider with completion status 10. Use case ends
2. Exception Flows	In all the cases below the relevant error code will be returned. Request could be rejected because of: 9. Invalid Service Provider ID or incomplete on-boarding - Please refer to the on-boarding information 10. Invalid Tax Agent or Customer ID 11. Unavailability of or internal error on the gateway services. Please contact IRD support 12. The request message is garbled /unreadable - IR sends the user a Signal Error Response Message. 13. There is an internal error within IR IT - IR IT uses responds

SUC Link Service Provider	
	to user with appropriate message. 14. The User name and/or password are incorrect - IR IT responds to user with appropriate message and records the access attempt. 15. The user is valid, however doesn't have the correct permissions to use this service - For example: Is not a Tax Agent. Or, User has had their access revoked due to abuse of the service. IR records the access attempt. 16. The structure of the information in the User Request is invalid. (i.e. does not conform to XML standard) - IT records failure and responds to user with appropriate message.
3. Alternatives	For initial transition of existing service provider consumers a bulk linking process will be used. The following use cases will be detailed in the Intermediation Build Pack when complete: • Unlink Service Provider • List Service Provider Links

6 Appendix B - Glossary

TODO – complete this section

Term	Meaning
Authentication	The process of verifying an identity claimed by or for a system entity. [RFC 2828]
Authorisation	A right or a permission that is granted to a system entity to access a system resource. [RFC 2828]
Build Pack	Details the technical requirements and specifications, processes and sample payloads for the specified activity
Business Processing	Processing by Inland Revenue systems in retrieving data and constructing the Payload (business information content) of a message.
Business Service	An integration interface (description) of the Solution which provides a set of business data and information in fulfilling the Service and is specified in this document. The Solution may offer more than one Business Service.
Confidential Information	Means, in relation to a party, any information (in any form whether written, electronic or otherwise): (a) relating to the business or operations of that party or its suppliers or customers; (b) disclosed by that party to the other party on the express basis that such information is confidential; or (c) which might reasonably be expected by that party to be confidential in nature;
Customer	A Customer is the party who is a tax payer or a participant in the social policy products that are operated by Inland Revenue. The Customer might be a person (an "individual") or a non-individual entity such as a company, trust, society etc. Practically all of the service interactions with Inland Revenue are about a Customer (e.g. their returns, accounts, entitlements etc.) even though these interactions might be undertaken by an Intermediary such as a tax agent on their behalf.
Credentials	Information used to authenticate identity, for instance an account username and password.
Data integrity	The property that data has not been changed, destroyed, or lost in an unauthorized or accidental manner. [RFC 2828]
Digital certificate	A certificate document in the form of a digital data object (a data object used by a computer) to which is appended a computed digital signature value that depends on the data object. [RFC 2828]
Digital signature	A value computed with a cryptographic algorithm and appended to a data object in such a way that any recipient of the data can use the signature to verify the data's origin and integrity. [RFC 2828]

Term	Meaning
ECDSA	In cryptography, the Elliptic Curve Digital Signature Algorithm (ECDSA) offers a variant of the Digital Signature Algorithm (DSA) which uses elliptic curve cryptography.
Encryption	Cryptographic transformation of data (called "plaintext") into a form (called "cipher text") that conceals the data's original meaning to prevent it from being known or used. If the transformation is reversible, the corresponding reversal process is called "decryption", which is a transformation that restores encrypted data to its original state. [RFC 2828]
GWS	Gateway Services—the name for the suite of web services that IR is providing.
HTML	Hypertext Markup Language.
HTTP	Hypertext Transfer Protocol is a networking protocol and is the foundation of data communication for the World Wide Web.
HTTPS	HTTP that uses SSL.
IAMS	Identity and Access Management—a logical component that performs authentication and authorisation. Physically it is a set of discrete hardware and software products, plug-ins and protocols. Usually implemented as separate External IAMS (XIAMS) and Internal IAMS.
IAS Build Pack	Identity and Access Build Pack
Intermediary	A party who interacts with Inland Revenue on behalf of a Customer. Inland Revenue's Customer is a Client of the Intermediary. There are several types of Intermediary including Tax Agents, PTSIs, PAYE Intermediaries etc.
Intermediation Service	
IP	Internet Protocol—the principal communication protocol in the Internet protocol suite for relaying datagrams across networks.
IR	Inland Revenue Department.
MEP	Message Exchange Pattern.
MSH	Messaging Service Handler.
Mutual authentication	Mutual authentication refers to two parties authenticating each other at the same time, being a default mode of authentication in some protocols (e.g. SSH) and optional in other (TLS)
Non-repudiation	Protection against false denial of involvement in a communication. [RFC 2828]
NSP	Inland Revenue's New Services Platform—includes START, XIAMS, the Application Publishing Service and supporting infrastructure.
NZISM	NZ Information Security Manual—the security standards and best practices for Government agencies. Maintained by the NZ Government Communications Security Bureau (GCSB).
OAuth 2.0	OAuth 2.0 is an industry-standard protocol for authorization
Operational	The commencement of Service.

Term	Meaning
Commencement	
Pattern	A constraint on data type values that require the string literal used in the data type's lexical space to match a specific pattern.
Payload	The business information content of the message and/or file(s) between IR and a Business Partner.
Service	The exchange, as enabled by the Solution, of information, data and/or funds for the purpose of Clients' tax administration by Tax Agents.
Service Provider Software	A Client Application is an operating instance of Software that is deployed in one or more sites. A number of deployment patterns are possible: 1. A single cloud based instance with multiple tenants and online users, 2. An on premise instance (e.g. an organisation's payroll system) 3. A desktop application with an online user. This is the computer software that contains interfaces to consume the services that Inland Revenue exposes. Software is developed and maintained by a Software Developer and subsequently deployed as one or more Client applications.
SFTP	Secure File Transport Protocol. SFTP 3.0 is used.
SOAP	Simple Object Access Protocol (SOAP) is a protocol specification for exchanging structured information in the implementation of Web Services in computer networks.
Solution	The technology components, systems and interface specifications constituting the Tax Agent Web Services capability which enables integration and communication across the Gateway channel between Inland Revenue and Tax Agents for the purpose of providing the Service.
Software Developer	The developer of a Tax Agent software package and its Gateway Channel integration capability which forms part of the Solution.
SSH	Secure Shell (SSH) is a cryptographic network protocol for operating network services securely over an unsecured network. Version 2.0 is used.
SSL	Secure Sockets Layer (SSL) is a cryptographic protocol that provides security for communications over networks such as the Internet.
START	Simplified Taxation and Revenue Technology—IR's new core tax processing application. It is an implementation of the GenTax product from FAST Enterprises.
System	The parts of the Solution operated by a single Business Partner; typically this term means the Business Partner's MSH.
Tax Agent	A tax agent who is formally registered as such with Inland Revenue.
TDS	Transaction Data Services
TLS1.2	Transport Layer Security version 1.2—the protocol that is observed between adjacent servers for encrypting the data that

Term	Meaning
	they exchange. Prior versions of TLS and all versions of SSL have been compromised and are superseded by TLS1.2.
URL	Universal Resource Locator—also known as a 'web address'.
User	The user referred to in this document is the user of the software provider accounting or tax package. This user needs delegated permissions on Customer tax accounts (potentially via a tax agency or other intermediary) in order to use TDS. The web logon used in eServices needs to be used in making IRD queries. This web logon must be granted permission there to access Customer Accounts
WSDL	Web Services Description Language (WSDL) is an XML-based language that provides a model for describing Web Services.
XML	EXtensible Markup Language

7 Appendix C—Document history

Version	Date	Description
0.1	15 Aug 2017	Internal only
0.2	31 Aug 2017	Draft for initial feedback
0.5	6 Nov 2017	Reworked draft for initial feedback