Inland Revenue
Te Tari Taake

Inland Revenue

Gateway Services Build Pack: Transactional Data Service (TDS) Bulk File Feed V0.80 2 February 2018

Date: 2 February 2018
Version: v0.80

UNCLASSIFIED

Inland Revenue
Te Tari Taake

Business Transformation

About this Document

This document is provided to Software 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 Software Providers need for implementing interfaces between their software and Inland Revenue TDS.

Document Control

Document Name	Gateway Services Build Pack - Transactional Data Service (TDS) - Bulk File Feed V0.80 2 February 2018
Document Version	0.80 Draft
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

Figures

Tables

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 Inland Revenue.

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

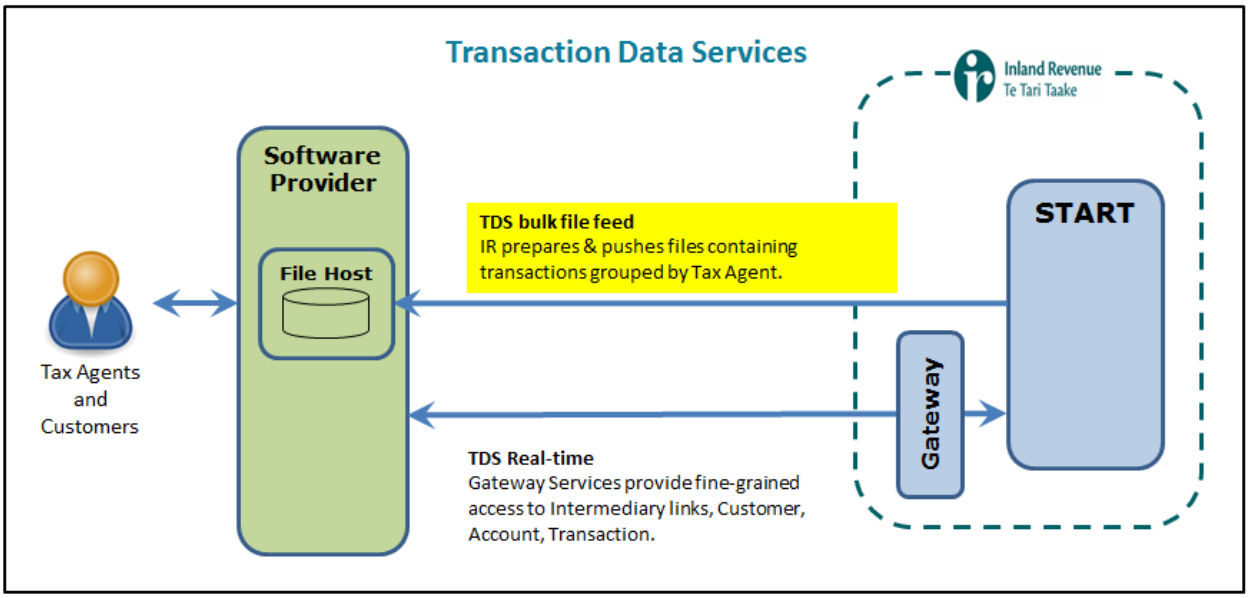

Figure 1: Transactional Data Services overview

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 Software 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 Inland Revenue's new system, START. See the TDS 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 Appendix A—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 Inland Revenue GitHub website here:

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

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

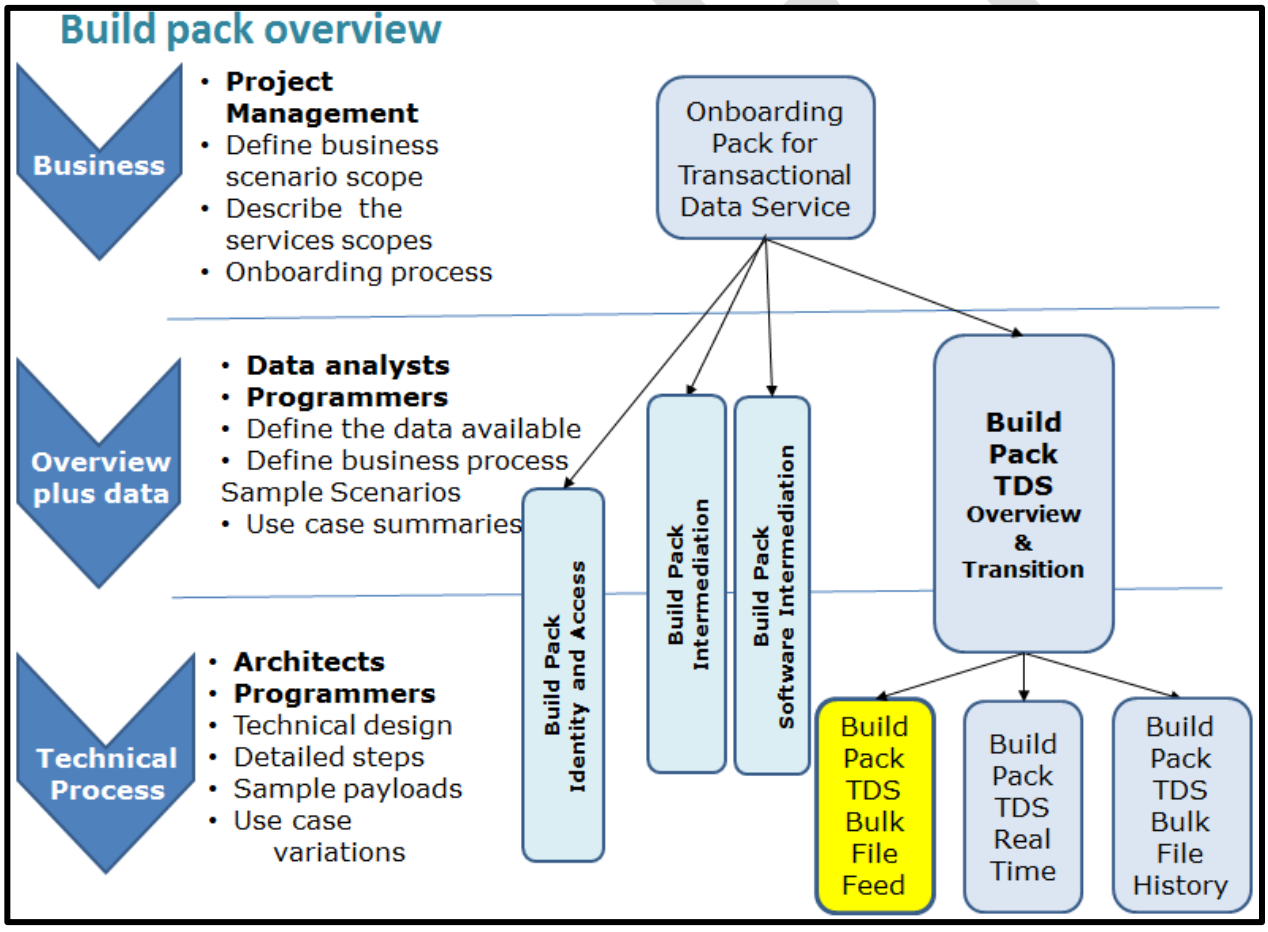

Figure 2: Onboarding and build pack structure for TDS

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 Software 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 Inland Revenue.
Software Intermediation Build Pack	Details the technical requirements and specifications for the linking of Tax Agents/Customers to Software Providers to enable these links to be used by the Bulk File Feed and Bulk file History Service.
Intermediation Build Pack	Details the technical requirements and specifications querying the links between Tax Agents and Clients to enable these links to be used by the TDS Real Time queries.

Table 1: Related documents

2 Technical design

2.1 Overview

The TDS Bulk File Feed is intended to be used by Software Providers where large quantities of data are required.

The TDS bulk file feed is based around a file transfer solution, where Inland Revenue will send (via SFTP) information to the Software 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.

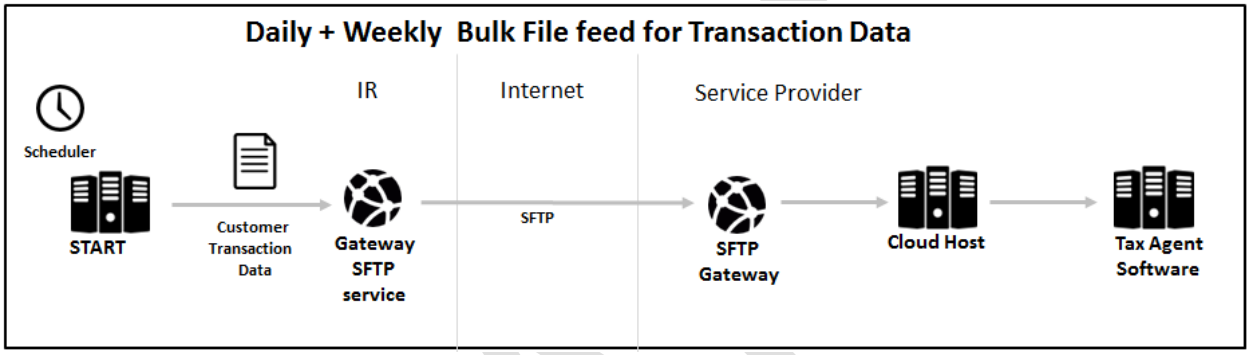

Figure 3: Daily and Weekly Bulk file feeds

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

Additionally Software Providers can request ad hoc manually generated files from Inland Revenue support. These files will be sent over the same SFTP channel when generated.

Each subscribing Software 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 Software Provider, a link needs to be established at Inland Revenue between the Tax Agent or Customer and the Software Provider – this link is maintained via the Software Intermediation Service. This is a new web service that is being made available to support the TDS Bulk File Feed.

Please see the TDS Overview Build Pack for the full process and context. More details about the Software Intermediation Service and the Tax Agent to client linking are available in the Software Intermediation Build Pack and 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 Software Provider linking are available in the Software Intermediation Build Pack.

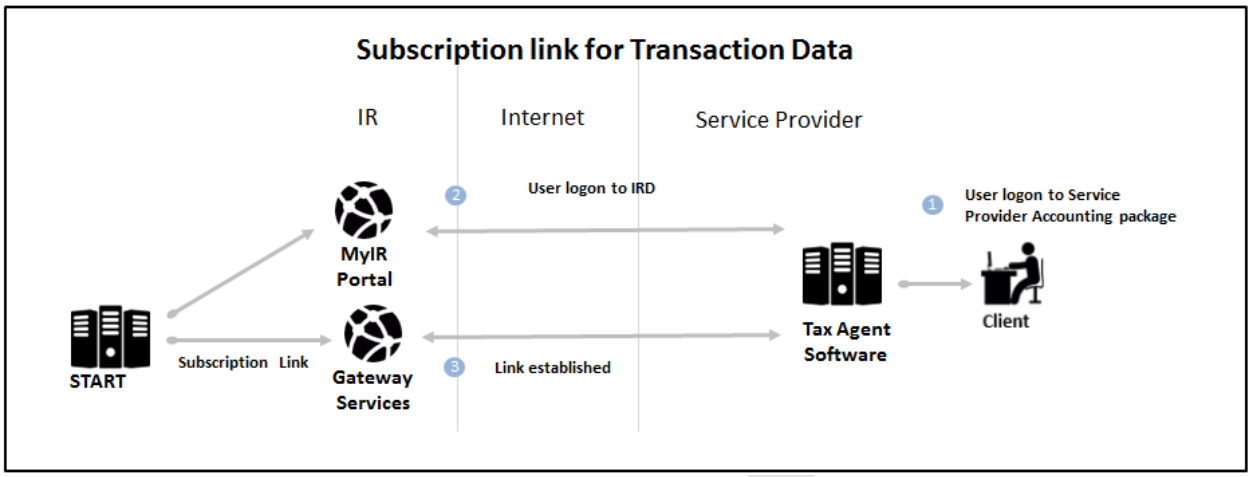

Figure 4: Subscription link for Transaction Data

The diagram above depicts the following steps:

1. The Tax Agent or other user logs into their Software Provider software application
2. The software will guide the user through a process of requesting daily bulk file updates including being directed to login to myIR portal (See Use case PUC210 Link Service Provider in the TDS Overview Build Pack.)
3. That login is then used by the software to establish a subscription link between the Software Provider and user (Tax Agent or direct Customer using their Software Provider software application)

2.2 Transfer mechanisms

2.2.1 Connectivity for bulk file feed

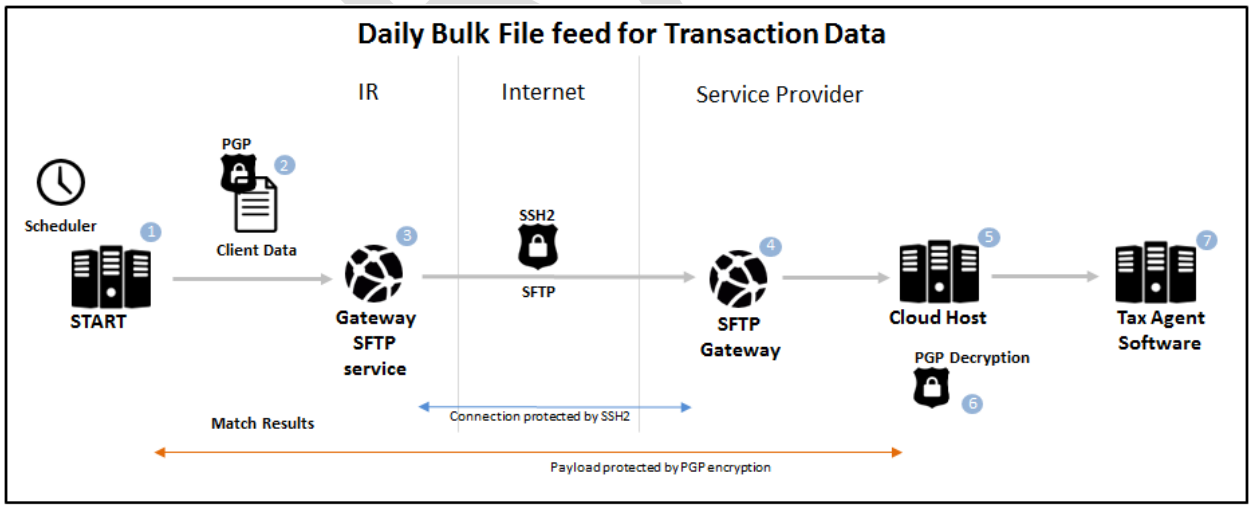

Figure 5: Daily feed - connectivity and security

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

Software Providers or parties planning to use the bulk file feed host an SFTP server that Inland Revenue 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 described in the Onboarding Section of the TDS Overview Build Pack.

Inland Revenue will provide its public key from a key pair to be set up for access to the Software Provider SFTP site. The exact keys and their nature will be agreed during the onboarding phase. For SFTP keys Inland Revenue will prefer to be NZISM compliant and therefore to use ECDSA keys. Where a Software Provider cannot support this RSA 2048 keys will be used. Inland Revenue 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. Inland Revenue thereby ensures that once a file is transferred to an endpoint only an authorised party can interpret it. As per PGP convention the receiver (Software Provider) keys are used by the sender (IR). 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:

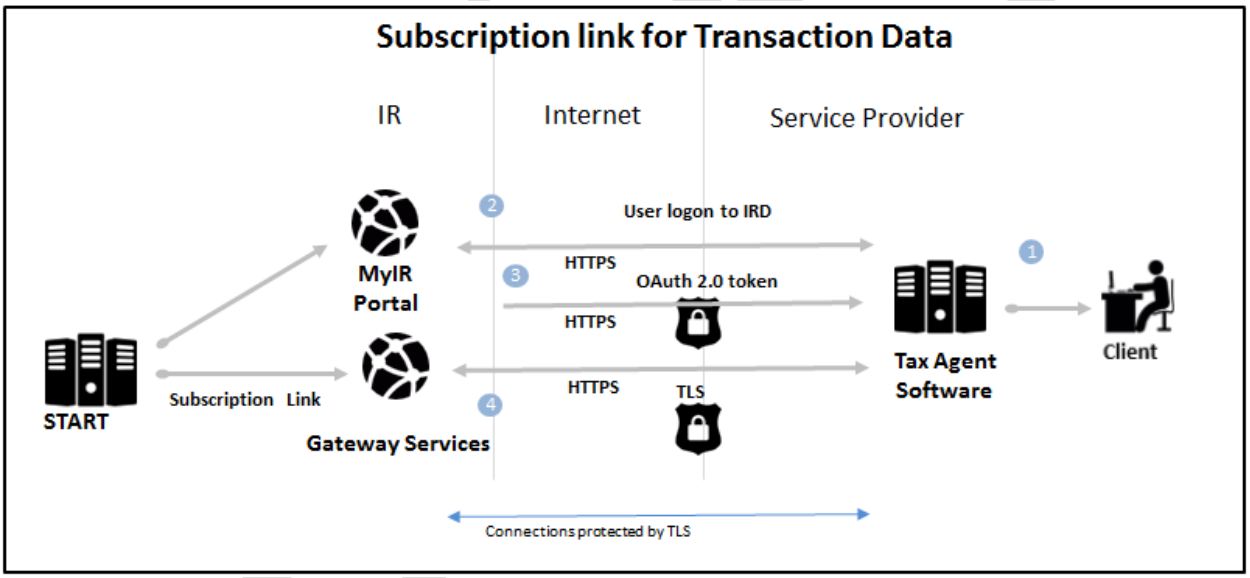

Figure 6: Subscription link for Bulk File Feed – connectivity and security

The numbers above refers to the following sequence of events:

1. Tax Agent or Customer uses their Software Provider software application
2. The user is presented with a browser showing a myIR logon the user needs to complete
3. Based on the myIR logon an OAuth token is eventually returned to the Software Provider software
4. That token and the Software Provider's connection to Inland Revenue are used to call the Software Intermediation Service to link the Tax Agent or Customer to the Software Provider so that any transaction data updates for the Tax Agent or Customer are sent to that Software 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 Software Providers might use it to validate that all intended files in the zip files are present, other Software 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 that content was received and unmodified.

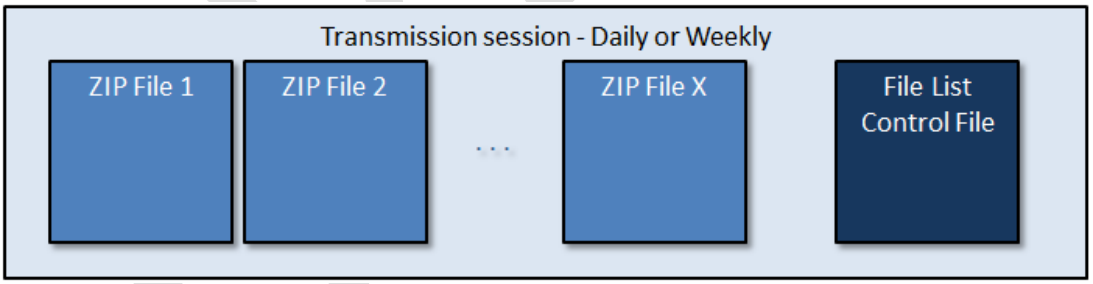

Figure 7: Zip Files plus Control file

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

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 a Software Intermediation link listing file. This file lists all the Tax Agents with a Software Intermediary link to the Software Provider.

2.3.2

Each Zip file will be PGP signed using the Inland Revenue key and PGP encrypted using the Software Providers Public PGP key.

There will initially be one Tax Agent file with the data for all the clients linked to the given Tax Agent. To optimise performance (of encryption processing) the files with Tax Agent information will be split across more than one file when they go beyond 100,000 transaction level elements.

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

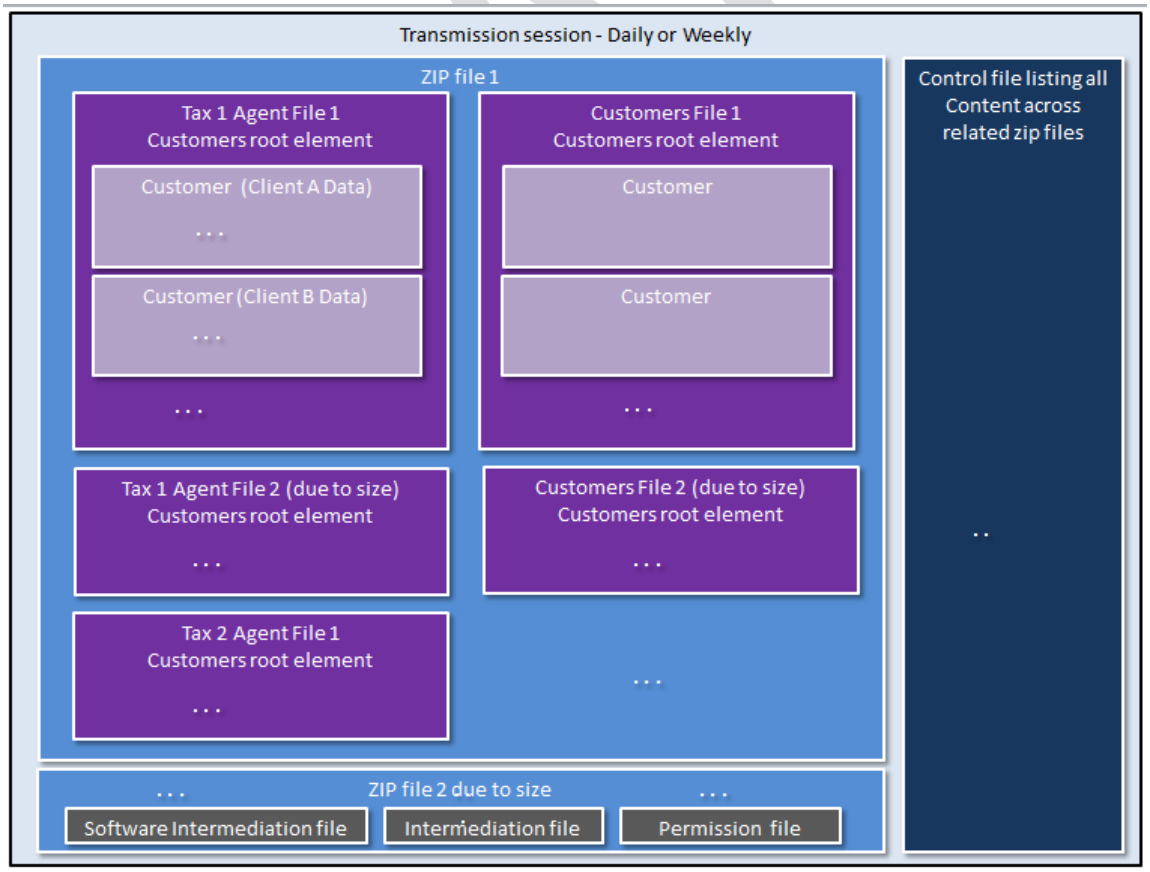

Figure 8: Zip files and their structure and summary control file

2.3.3 File structure and content and variations

Where the Software Provider linked user is not a Tax Agent or intermediary but a direct Customer, there will be a single file containing all this data across all Customers.

Where the Software 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:

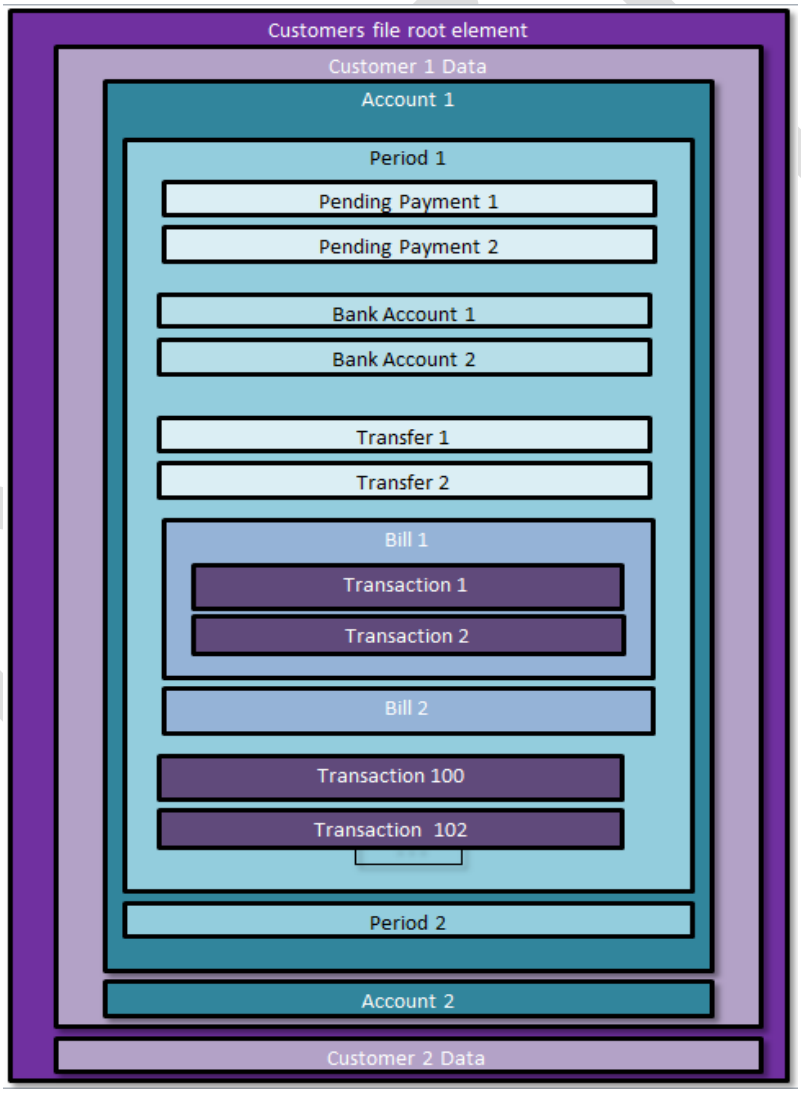

Figure 9: File content

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 Software 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 Software 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 Software 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) Software 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.

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.

Any ambiguity this might have created as to whether the agent or their clients are still linked is removed by the addition of the software intermediation link listing file and intermediation link listing file. These files will only appear in the last zip file for a file set. They will also be listed in the control file.

The filenames of all these files will indicate when did the batch run generate the file – as at what time does this file set provide all known transactions and intermediation links?

2.4 Content of Software Intermediation link file

The last zip file will also contain the intermediation link file.

This file summarises

- Which Tax Agents are linked to this Software Provider.
- Which Customers are linked directly to this Software Provider as users

2.5 Content of Intermediation link file

The last zip file will also contain the intermediation link file.

This file summarises

- Which Customer accounts are linked to the Tax Agents that are linked to the Software Provider

2.6 Content of Permissions file

The last zip file will also contain the permissions file.

This file summarises

- The names of the myIR logons associated with linked Tax Agents and their roles
- The subset of client lists visible to each of those
- The names of the myIR logons associated with linked Customer Accounts and their roles
- The subset of Customer Accounts visible to each of those

2.7 Transfer of files – processing

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

1. Receive control file to help calculate processing
2. Receive one or multiple zip files from Inland Revenue
3. Validate file list in the control file matches the list of zip files received, if not, escalate to Inland Revenue 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 Inland Revenue 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

2.8 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 Inland Revenue 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 previously 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.

2.9 File naming conventions

2.9.1 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_<software_platform_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
<software_platform_id>	ID allocated to Software Platform by Inland Revenue during the onboarding process	Numeric 10 digit, should remain constant for a given Software Provider
<file sequence>	Next number after zip file count – last number in sequence	1 2 3 …
<timestamp>	Time file was created yyyyMMddHHmmssffff	e.g. 201710100921548813
<environment>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD

Table 2: Naming of Control File

2.9.2 ZIP files names

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

Format:
<Frequency>_PROVIDER_<software_platform_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
<software_platform_id>	ID allocated to Software Platform by Inland Revenue during the onboarding process	Numeric 10 digit, should remain constant for a given Software 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>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD

Table 3 : Naming of Zip files

The following file types are embedded in the zip files:

2.9.2.1 Tax agent files

Format:
<Frequency>_AGENT_<tax_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	AGENT

Part	Format	Possible values
	to a Tax Agent or other Intermediary	
<tax_agent_id>	ID allocated to Tax Agent by Inland Revenue	Numeric 9 digit IRD Number of Agent
<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. 20171010092154813
<environment>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD – Inland Revenue Production

Table 4 : Naming of Tax Agent files

2.9.2.2 Customer file

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

For example:
DAILY_CUSTOMER_000000000_0_20171011532239353_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 the Software Provider	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. 20171010092154813

<environment>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD

Table 5 : Naming of Customer Files

2.9.2.3 Software Intermediation file

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

Format:
<Frequency>_PROVIDER_<software_platform_id>_<filesequence>_<timestamp>_<environment>_SOFTWARE_INTERMEDIATION.xml

For example:
DAILY_PROVIDER_1500011034_0_201710100921548813_NZD_SOFTWARE_INTERMEDIATION.xml

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
PROVIDER	Constant	PROVIDER
<software_platform_id>	ID allocated to Software Platform by Inland Revenue during the onboarding process	Numeric 10 digit, should remain constant for a given Software 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>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD
SOFTWARE_INTERMEDIATION	Constant	SOFTWARE_INTERMEDIATION

Table 6 : Naming of Software Intermediation Files

2.9.2.4 Intermediation file

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

Format:
<Frequency>_PROVIDER_<software_platform_id>_<filesequence>_<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
<software_platform_id>	ID allocated to Software Platform by Inland Revenue during the onboarding process	Numeric 10 digit, should remain constant for a given Software 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>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD
INTERMEDIATION	Constant	INTERMEDIATION

Table 7 : Naming of Intermediation Files

2.9.2.5 Permissions 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>_<filesequence>_<timestamp>_<environment>_PERMISSION.xml

For example:
DAILY_PROVIDER_1500011034_0_201710100921548813_NZD_PERMISSION.xml

Part	Format	Possible values
<Frequency>	See list:	DAILY WEEKLY MANUAL
PROVIDER	Constant	PROVIDER
<software_platform_id>	ID allocated to Software Platform by Inland Revenue during the onboarding process	Numeric 10 digit, should remain constant for a given Software 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. 20171010092154813
<environment>	Inland Revenue environment	Three letters Production: PRD Partner testing: XZS XZT Internal IRD testing: NZT NZD
PERMISSION	Constant	PERMISSION

Table 8 : Naming of Permissions Files

2.10 Sample payloads

This service has no incoming requests so what follows is details of the file contents for files which are sent.

Please refer to the Software Intermediation and Identity and Access build packs for related requests.

The picture below is an extract of what files can contain; please see the attached sample files below for the full context. For descriptions of the nature of the data please refer to the TDS Overview Build Pack.

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

Sample Customer File

```
<?xml version="1.0"?>
- <Customers xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema">
    - <Customer IDType="IRD" ID="123041507">
        - <ClientList ListIDType="" ListID="">
            - <Account ID="123-041-507-GST002" MaxActivity="1900-01-01T00:00:00" Balance="0.00" Cease="9999-
              12-31" Commence="2017-10-31" FilingFrequency="GST2M1" Type="GST">
                <Period Balance="0.00" FilingFrequency="" Activity="1900-01-01T00:00:00" PenaltyForecast="0.00"
                  FilingPeriod="9999-12-31" Begin="9999-12-31" End="9999-12-31" Penalty="0.00" Tax="0.00"
                  Interest="0.00" InterestForecast="0.00" Other="0.00" OtherForecast="0.00" Credit="0.00"
                  CreditForecast="0.00"/>
              </Account>
          </ClientList>
      </Customer>
    - <Customer IDType="IRD" ID="123068629">
        - <ClientList ListIDType="" ListID="">
            - <Account ID="123-068-629-AIL002" MaxActivity="2017-09-15T08:48:18" Balance="0.00" Cease="9999-
              12-31" Commence="2017-04-01" FilingFrequency="AILMO" Type="AIL">
                <Period Balance="0.00" FilingFrequency="AILMO" Activity="9999-12-31T00:00:00"
                  PenaltyForecast="0.00" FilingPeriod="2017-04-30" Begin="2017-04-01" End="2017-04-30"
                  Penalty="0.00" Tax="0.00" Interest="0.00" InterestForecast="0.00" Other="0.00"
                  OtherForecast="0.00" Credit="0.00" CreditForecast="0.00"/>
                <Period Balance="0.00" FilingFrequency="AILMO" Activity="9999-12-31T00:00:00"
                  PenaltyForecast="0.00" FilingPeriod="2017-05-31" Begin="2017-05-01" End="2017-05-31"
                  Penalty="0.00" Tax="0.00" Interest="0.00" InterestForecast="0.00" Other="0.00"
                  OtherForecast="0.00" Credit="0.00" CreditForecast="0.00"/>
                <Period Balance="0.00" FilingFrequency="AILMO" Activity="9999-12-31T00:00:00"
                  PenaltyForecast="0.00" FilingPeriod="2017-06-30" Begin="2017-06-01" End="2017-06-30"
                  Penalty="0.00" Tax="0.00" Interest="0.00" InterestForecast="0.00" Other="0.00"
                  OtherForecast="0.00" Credit="0.00" CreditForecast="0.00"/>
                <Period Balance="0.00" FilingFrequency="AILMO" Activity="9999-12-31T00:00:00"
                  PenaltyForecast="0.00" FilingPeriod="2017-07-31" Begin="2017-07-01" End="2017-07-31"
                  Penalty="0.00" Tax="0.00" Interest="0.00" InterestForecast="0.00" Other="0.00"
                  OtherForecast="0.00" Credit="0.00" CreditForecast="0.00"/>
              - <Period Balance="0.00" FilingFrequency="AILMO" Activity="2017-09-15T08:48:18"
                PenaltyForecast="0.00" FilingPeriod="2017-08-31" Begin="2017-08-01" End="2017-08-31"
                Penalty="0.00" Tax="200.00" Interest="0.00" InterestForecast="0.00" Other="0.00" OtherForecast="0.00"
                Credit="-200.00" CreditForecast="0.00">
                  - <Bill Balance="0.00" PenaltyForecast="0.00" Penalty="0.00" Tax="200.00" Interest="0.00"
                  InterestForecast="0.00" Other="0.00" OtherForecast="0.00" Credit="-200.00" CreditForecast="0.00"
                  Due="2017-09-20" BillIDForDisplay="1" BillID="582555648">
                      <Tx Type="RTNORI" LinkID="0" Effect="2017-09-01" Post="2017-09-14T07:37:01.783"
                        TransID="1821848064" Amt="200.00"/>
                  </Bill>
                  <Tx Type="PYMRTN" LinkID="0" Effect="2017-09-13" Post="2017-09-15T08:48:17.983"
                    TransID="826161152" Amt="-200.00"/>
                </Period>
                <Period Balance="0.00" FilingFrequency="AILMO" Activity="9999-12-31T00:00:00"
                  PenaltyForecast="0.00" FilingPeriod="2017-09-30" Begin="2017-09-01" End="2017-09-30"
                  Penalty="0.00" Tax="0.00" Interest="0.00" InterestForecast="0.00" Other="0.00"
                  OtherForecast="0.00" Credit="0.00" CreditForecast="0.00"/>
              </Account>
          </ClientList>
      </Customer>
</Customers>
```

Figure 10 : Sample Customer file

Sample AGENT file

```
<?xml version="1.0"?>
- <Customers xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  - <Customer IDType="IRD" ID="123025148">
    - <ClientList ListIDType="ACCIRD" ListID="123025148">
      - <Account ID="123-025-148-GST004" MaxActivity="2018-01-
        09T14:44:31.303" Balance="678.20" Cease="9999-12-31"
        Commence="2017-01-01" FilingFrequency="GST2M1" Type="GST">
        - <Period Balance="678.20" FilingFrequency="GST2M1" Activity="2018-01-
          09T14:44:31.303" PenaltyForecast="0.00" FilingPeriod="2017-07-31"
          Begin="2017-06-01" End="2017-07-31" Penalty="284.10"
          Tax="376.12" Interest="17.98" InterestForecast="0.00" Other="0.00"
          OtherForecast="0.00" Credit="0.00" CreditForecast="0.00">
          - <Bill Balance="2.54" PenaltyForecast="0.00" Penalty="2.54"
            Tax="0.00" Interest="0.00" InterestForecast="0.00" Other="0.00"
            OtherForecast="0.00" Credit="0.00" CreditForecast="0.00"
            Due="2018-01-09" BillID="1585924096">
              <Tx Type="PNLWSD" LinkID="0" Effect="2018-01-09"
                Post="2018-01-09T14:44:30.007" TransID="1808943104"
                Amt="2.54"/>
            </Bill>
          </Period>
        </Account>
      </ClientList>
    </Customer>
</Customers>
```

Figure 11: Sample Agent file

2.10.3 Sample Software Intermediation file

```
<?xml version="1.0"?>
- <SoftwareIntermediationSummary RetrievedDate="2018-01-30T18:05:16.407"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
    <CustomerLink AccountType="GST" IDType="IRD" ID="123041507"/>
    <CustomerLink AccountType="AIL" IDType="IRD" ID="123068629"/>
    <AgentLink IDType="IRD" ID="123105028"/>
</SoftwareIntermediationSummary>
```

Figure 12: Sample Software Intermediation File

Sample Intermediation file

```
<?xml version="1.0"?>
- <IntermediationSummary RetrievedDate="2018-01-30T18:05:16.407"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  - <AgencyLinks IDType="IRD" ID="123105028">
    - <ClientList IDType="IRD" ID="123105028">
        <Link IDType="IRD" ID="123024052" AccountType="IPS"/>
        <Link IDType="IRD" ID="123025148" AccountType="GST"/>
        <Link IDType="IRD" ID="123041028" AccountType="DWT"/>
        <Link IDType="IRD" ID="123051106" AccountType="FBT"/>
        <Link IDType="IRD" ID="123065883" AccountType="GMD"/>
        <Link IDType="IRD" ID="123068629" AccountType="AIL"/>
        <Link IDType="IRD" ID="123088522" AccountType="GST"/>
        <Link IDType="IRD" ID="123088883" AccountType="DWT"/>
      </ClientList>
    </AgencyLinks>
</IntermediationSummary>
```

2.1

Figure 13: Sample Intermediation File

2.11 Schema

File Type	Reference
Basic Simple types	 Common.v1.xsd
Structure	 TransactionBulkList_ v0.2.xsd

Table 9 : Schema

2.12 Samples

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

File type	Reference
Zip and control files	 2018-01-30.zip

Table 10 : Samples

3 Onboarding

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

3.1 Software 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 Inland Revenue information

- All Inland Revenue'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 TDS Overview Build Pack for a more detailed explanation of the stages of TDS usage, summarised below.

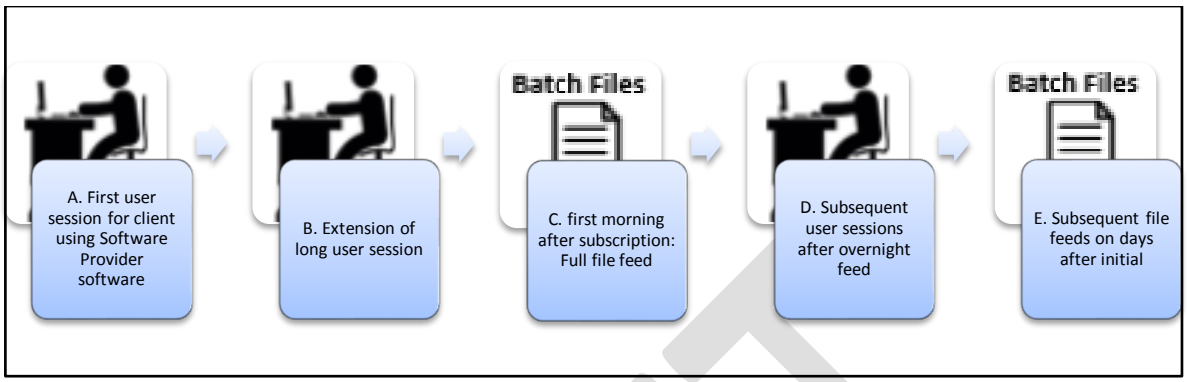

Figure 14: Overall process

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

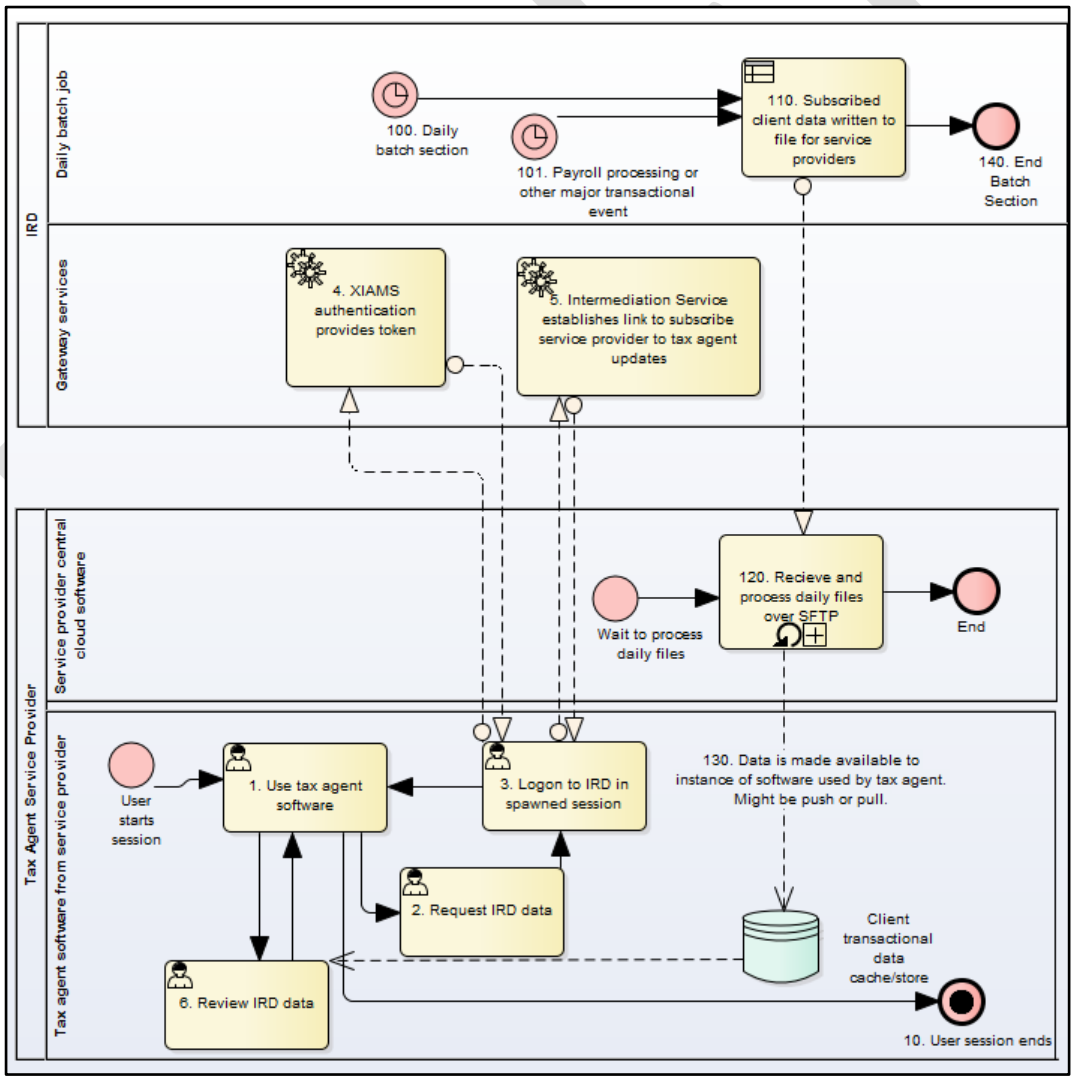

Figure 15: Process model for TDS Bulk File Feed

A. Software Provider and Inland Revenue setup SFTP path and exchange credentials during onboarding

1. Prerequisite (not shown on process model). See the Onboarding section of the TDS Overview Build Pack for more information

B. User session for agent or customer using Software Provider software

1. User signs onto Tax Agent software from Software Provider
2. User decides to request access to some Inland Revenue tax data
3. Software Provider software lets user starts an Inland Revenue session by logging onto the Inland Revenue 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 Software Provider software then uses this token in a call to the Software Intermediation Gateway service to request a link between the Software Provider software and the user (whether Tax Agent or direct Customer). This link subscribes the Software 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 Inland Revenue overnight processing runs where most updates are done, towards the end of the sequence a job kicks off to update subscribed Software 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 Software Providers their data gets written to file and aggregated to send to the Software Providers.
120. The Software 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 Software 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 Software Provider software
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 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 TDS Overview Build Pack the transaction data services provide data into an overall business process that is controlled by the Software 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 Software 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 Inland Revenue 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.

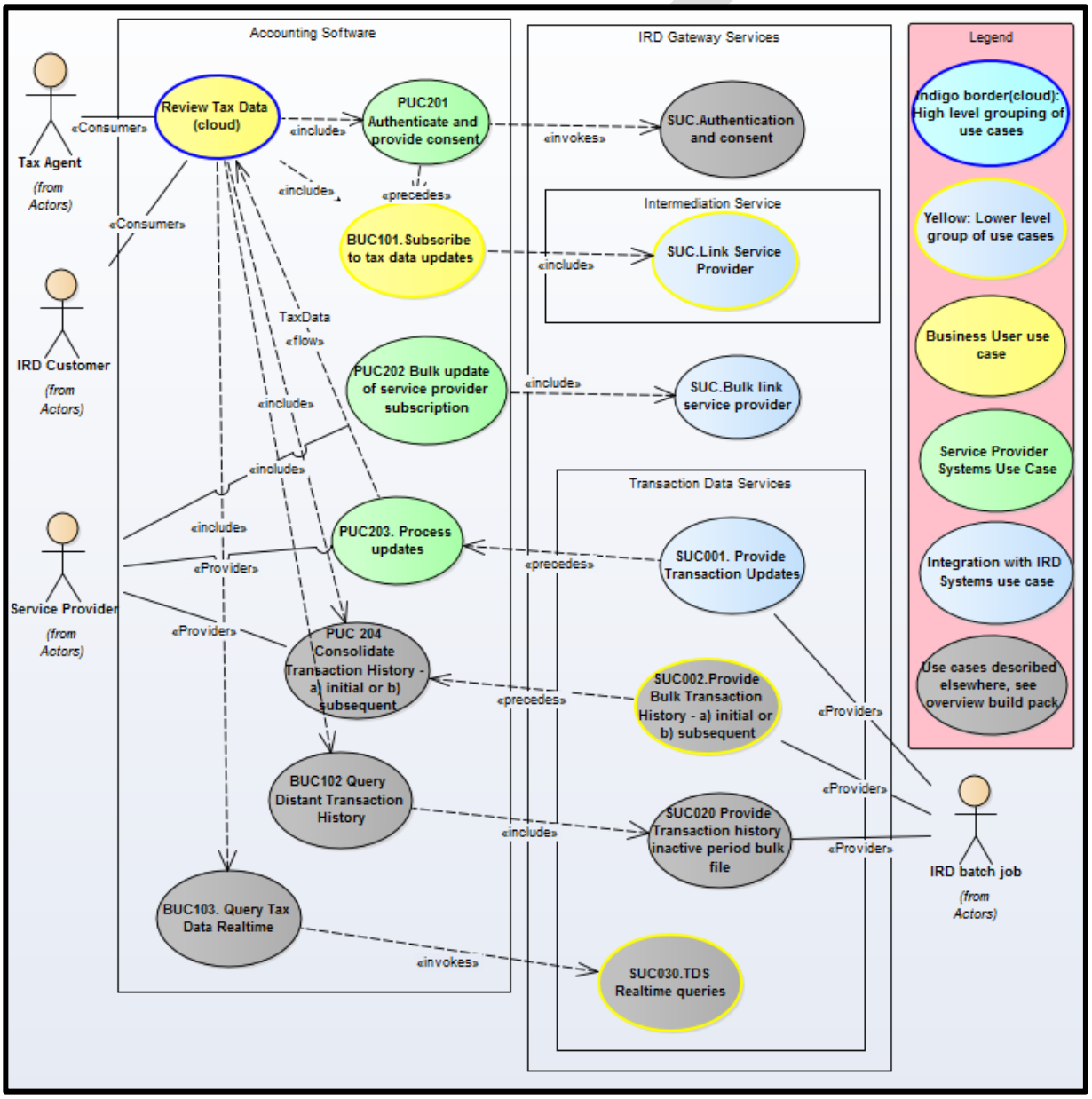

Figure 16: Use case overview

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
Software Provider	Some systems use cases on the Software 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 Software 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 Software Provider	Software Intermediation BP
			SUC Bulk Link Software Provider to seed the Inland Revenue 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 11: 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	Inland Revenue Transaction Data Services
Secondary Actor	Software Provider Software
Description	The use case goal is to retrieve data for all Tax Agents/Customers linked to the Software Provider Software, format and package it as described in chapters above and send via SFTP to the Software Provider central location
Inland Revenue systems	START
Pre-Conditions	Software Provider Software is onboarded and therefore authorised to communicate with TDS (See the TDS Onboarding Build Pack) A link / subscription has been established between the Software Provider and Tax Agent and / or Customer Account identifying which entities' tax data to send to the Software Provider and how to group it. (See the TDS Overview Build Pack)
Triggers	Inland Revenue 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 Inland Revenue 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 Software 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 Software Provider Software and PUC203 can be completed by the Software Provider. (See TDS Overview Build Pack)
Use Case Scenarios	
1. Normal Flow	Inland Revenue daily batch processing for Software Provider subscription generates PGP signed and encrypted ZIP files. These files are sent to the Software Provider at a Software 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 Software Provider is receiving from Inland Revenue daily. 1. For each of Daily / Weekly / Manual cycle as applicable: 2. Iterate through on-boarded Software Providers and their linked Customers as well as their linked Tax Agents and

Use Case SUC001 Provide Transaction Updates	
	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 Software Provider Software central location via SFTP as described above in this document 6. This system use case ends Software Provider use case PUC203 then completes at each Software Provider The steps after the SFTP transfer completes depend on the Software Provider Software. Once all the file content has been transferred the Software 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 Inland Revenue plans a maintenance outage and will not be able to do a daily batch run Software Providers will be notified through agreed communication channels established during onboarding If Inland Revenue is unable to connect to the Software Provider SFTP site the system will escalate for the Inland Revenue support team to contact the Software Provider and agree a plan to restore connectivity and reschedule transmission If SFTP transmission fails it escalate to the Inland Revenue support team to contact the Software 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 Software Provider finds there are issues like file corruption they will contact Inland Revenue 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 Software 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 Software Provider and Inland Revenue support to not generate volumes that degrades the overall experience for Software Provider users or those connecting to Inland Revenue.

Table 12: Use Case SUC001

5 Appendix A—Glossary

Term	Meaning
Authentication	The process of verifying an identity claimed by or for a system entity. [RFC 2828]
Build Pack	Details the technical requirements and specifications, processes and sample payloads for the specified activity
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.
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.
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 Inland Revenue is providing.
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	The Intermediation Service is a new Gateway Service for creating and maintaining delegated access relationships between intermediaries and their clients. These relationships enable access by the intermediary to a resource (e.g. an account, correspondence etc.) that belongs to their client. There are several types of intermediaries such as Tax Agents, book keepers, PAYE Intermediaries.
IP	Internet Protocol—the principal communication protocol in the Internet protocol suite for relaying datagrams across networks.
MSH	Messaging Service Handler.

Term	Meaning
NZISM	NZ Information Security Manual—the security standards and best practices for Government agencies. Maintained by the NZ Government Communications Security Bureau (GCSB).
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 Inland Revenue 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.
Software Provider Software	A Client Application is an operating instance of Software Provider 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.
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.
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—Inland Revenue'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 they exchange. Prior versions of TLS and all versions of SSL have

Term	Meaning
	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 myIR logon used in eServices needs to be used in making Inland Revenue queries. This myIR logon must be granted permission there to access Customer Accounts.
XIAMS	External IAMS—an instance of IAMS that authenticates and authorises access by external parties, i.e. customers, trading partners etc. as opposed to internal parties such as staff.
WSDL	Web Services Description Language (WSDL) is an XML-based language that provides a model for describing Web Services.
XML	EXtensible Markup Language

6 Appendix B—Document history

Version	Date	Description
0.8	31/01/2018	Small cosmetic changes – spelling and caps - no change to sense Removal of Appendix A – now included in Intermediation Build Pack Formatting changes Update for terminology such as Software Provider, distinction between Software Intermediation and Intermediation, Software Platform ID, Environments Update for Tables and Figures New Build Pack Overview diagram Update for Links to GitHub
0.5	06/11/2017	Reworked draft for initial feedback
0.2	31/08/2017	Draft for initial feedback
0.1	15/08/2017	Internal only