



Client Communication API design discussion key outcomes

OFFICIAL
External

| | | | |
|--------------------|--|----------------------|-----------------------|
| Title: | Client Communication API design discussion | | |
| Issue date: | 19 June 2020 | | |
| Venue: | Webinar | | |
| Event date: | 25 May 2020 | Start: 2:00pm | Finish: 3:00pm |

| | | | |
|---------------------|-----------------|-----------------------|--------------|
| Facilitator: | Sonia Lark | | |
| Contact: | Mikaela Jenkins | Contact phone: | 02 9734 1014 |

| | |
|-------------------------------------|--|
| Attendees: names/section | <p>ATO:</p> <p>David Baker – Director, Intermediaries Digital Services Peter Evans – Director, Digital Communications & Identity Services Donna Duncan – Director, Digital Communications & Identity Services Brendan Kee – Digital Wholesale Integration Services Damien Choy – Application Architecture and Design Sonia Lark – Digital Partnership Office Kylie Johnston – Director, Digital Partnership Office Mikaela Jenkins - Digital Partnership Office Danielle Miller - Digital Partnership Office</p> <p>Industry:</p> <p>Ron Drost – Digital Disruption Solutions Simon Smart – Etax</p> |
|-------------------------------------|--|

Albert Lilie – HostOne
Sandeep Gopalan - GovReports
Phil Martin – KPMG
Mike Behling – MYOB
Kevin Johnson – Reckon
Karl Farrand – TaxLab
Vikas Kumar – Wolters Kluwer
Anthony Migliardi - Xero

**Apologies:
name/section**

ATO:

Tyson Andrews – Director, Digital Communications & Identity Services
Craig Hughes – Digital Wholesale Integration Services
Danny Figueiredo - Digital Partnership Office

Industry:

Mike Denniss – Class
Kevin Zang – Class
Sha Sha - Class
Nathan Kerr – HostOne
Nic Taylor – HostOne
Lisa Miks – MYOB
Darin Carter - Sage
Danna Zheng – Wolters Kluwer

Agenda item: 1 – Welcome and recap of workshop insights – Sonia Lark

Sonia welcomed the group and thanked them all for attending. She explained work has continued in the background even during COVID-19 on the Client Communication API design. The outcomes from the workshop in March were issued today in the meeting invite and will be published in the next couple of days as publishing has now resumed. Please let the DPO know if you require a copy of the pack issued in March, which was run through on day two of the workshop.

Sonia introduced attendees from the ATO and explained the purpose of today's discussion is to cover CCL. Preferencing will not be a focus but we can take questions on notice. During the March workshop, the group discussed MVP and agreed on what was wanted in the Client Communication service. Today we will run through what the service will look like with Damien discussing the current design and Brendan covering the MST design.

Agenda item: 2 – Design for document list and retrieval for wholesale consumption – Damien Choy

Damien explained guidance is needed from the group to assist in finalising the design. He began with a recap from the last meeting, identifying three main CCL APIs:

1. List
2. Get single
3. Get batch

Documents are currently all PDFs, which can be moved to data payloads or XML at a later stage.

Use cases have been considered and the main concern is the potential loads. Damien provided an overview of the different systems which must be navigated to provide the service. This included the SBR entry point, which is managed by Brendan Kee's area, the next is infrastructure and finally repository. It was confirmed the backend is shared with ATO online and retail users and separation is required to prevent usage of the CCL impacting online users and vice versa. Batching and schedules will be implemented to manage the loads.

Use cases outlined:

1. List might be used daily for new items of communication
2. Get single would be used for ad-hoc requests and high usage was not expected
3. Get batch had the most interest and would be used for high volumes.

Alternative mechanisms for foundation loads are being explored due to large volumes. Get batch would be appropriate for daily retrieval rather than large historical loads.

Agenda item: 3 – Message design for Client Communication API – Brendan Kee

Brendan advised we are looking at the provision of communications record, specifically main value pairs rather than well-defined elements. As we extend the service, we want to ensure that we don't create breaking changes and force an update. Currently, searches can be done by date etc and up to ten individual entities can be specified.

The ABN / TFN fields are mandatory and all other fields are optional. We may need to implement rules to avoid high volume requests due to internal load restrictions such as searching by limited date ranges.

PDFs will be delivered in delimited fashion, like batches. We are defining the delimited structure – it is well defined on the way in, but the response needs to be updated.

Agenda item: 4 – Group review and discussion – All

Damien highlighted the key concern is size of the list / retrievals. If retrievals are too high during peak times, it will impact other services. Polling two or three times a day is acceptable but higher volumes can cause issues. As such, we are looking at where limits / caps can be implemented to avoid excessive use. For example, after pulling foundational data (which can be actioned over several days to reduce loads), can list requests be limited to three business days or by client ID, document type etc?

The group noted highest volumes would occur during foundational loads and when DSPs obtain a new client. One estimate of service usage was 10,000 documents per day and 70,000 per day during peak. Tax agents obtaining a new client would be less of a burden as foundational data would only be a few hundred documents. Once foundational data was pulled, only the delta would be needed. It was noted developers may exploit maximum limits and find a way around the caps, but they could be excluded from the system if required.

Processing times

DSPs wanted to understand the batch bulk queue and processing times for obtaining documents from the service. Damien advised the items are not stored in the mainframe and the solution is predominantly in midrange. It is the same service supporting ATO online today and users are not waiting minutes for documents.

Get single is hosted on SRP while List batch on BBRP – alternative hosting methods are being investigated for Get bulk. Loads over both channels are a concern. If all traffic was over SRP, it would harm the backend. Bulk data still needs to be stored, as aggregation is needed prior to retrieval.

List single results will be capped to 100. Where there are more than 100 results, messaging will be provided to advise users to either refine their search or use BBRP. Batch will be unrestricted and can be used where a greater number of results is needed.

The response time for List batch is estimated as half an hour, but the time is dependent on load. Better guidance will be provided after performance testing.

Past-dated communication will be captured by the service as current design thinking uses the date that the document was written to CCL, rather than when it was created. As long as developers query what communication is new, the documents will be retrieved.

Reducing loads

The group was flexible about setting limits on the service such as restricting by date range and client ID. Another option is to request high volumes during out of hours and the group will work on appropriate windows. During midnight to 5:00am is considered optimal for polling and retrievals as it is after large batching by the ATO, which usually ends by 10pm

Agenda item: 4 – Group review and discussion – All

but can stretch longer during tax time etc. NOA and SOA are not generated early morning. SMS and emails are created during the day by frontline staff though there can be lulls around midday and close of business.

DSPs wish to be engaged on the alternative mechanism for pulling foundational data. The bulk service may take place over several days. Perhaps an initial list could be provided then cut by priority. The system still needs to be protected while the service is live. As the channel is shared by everybody, the system cannot be overloaded. S3 was suggested in preference to SFTP.

For DSPs obtaining new clients, whole of practice data will not be pulled through the service. There will be a listing by TAN and may trickle-feed documents to avoid high volumes.

Security

Damien confirmed that PDF documents will not have TFNs hashed. The DPO will consider whether passing the documents onto a third-party, such as a document management service is permitted within the operational framework. Kylie Johnston agreed to take the use case offline to discuss.

Timeline

A definite timeframe for delivering the service cannot be provide at this stage. There are concerns about load which have changed the timeline, as well as COVID-19. We are working on alternatives and will performance test the service. We will reconvene shortly to run through timeframes as well as pulling foundational data.

Agenda item: 5 – Wrap up and next steps – Sonia Lark

Sonia confirmed that the group was in agreeance on the design, apart from the mechanism for the initial pulls. Damien outlined the next steps are to nail down caps on usage, a potential alternative mechanism for file transfer and validation. June is a realistic time to reconvene as we are working on the build and dependency.

Sonia asked the group to reach out if they have anything further to raise, which can be passed on to Damien and the technical team.