

Technical Design Document

**Template version 2.0**

Operations Systems

DSA Entitlement Database

Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Revision date | Author | Version No. | Details/Comments |
| 2021/03/01 | Andy Neve | 1.0 | Initial Version |
| 2021/10/08 | Andrew Mackenzie | 1.1 | Additional information |
| 2021/10/13 | Andrew Mackenzie | 1.2 | Technical Diagrams and Data Dictionaries added |
| 2021/11/03 | Andrew Mackenzie & Andy Neve | 1.3 | Further detail provided |
| 2021/11/17 | Rahul Joshi | 1.4 | Removed unwanted content, added High level diagrams, updated Incomplete information |

References

|  |  |  |  |
| --- | --- | --- | --- |
| # | Title | Link | |
| 1 | Solution in principle | |  |
| 2 | TDD related | |  |
| 3 | Project plan | |  |

Contents

[2 Solution Background 3](#_Toc88048892)

[2.1 Solution Context 3](#_Toc88048893)

[2.2 Objectives 3](#_Toc88048894)

[2.3 Assumptions 4](#_Toc88048895)

[2.4 Risks 4](#_Toc88048896)

[2.5 Issues 5](#_Toc88048897)

[2.6 Dependencies 5](#_Toc88048898)

[2.7 Out of Scope 6](#_Toc88048899)

[2.8 High-Level Delivery plan 6](#_Toc88048900)

[3 Requirements and Characteristics 6](#_Toc88048901)

[3.1 Business Processes Supported 6](#_Toc88048902)

[3.2 High-Level Functional Requirements 7](#_Toc88048903)

[3.2.1 Business logic 8](#_Toc88048904)

[3.2.2 Integration 9](#_Toc88048905)

[3.3 Information Architecture 9](#_Toc88048906)

[3.3.1 Information Context 9](#_Toc88048907)

[3.3.2 Information Entities 9](#_Toc88048908)

[3.3.3 High Level Workflows 10](#_Toc88048909)

[3.4 User community description 12](#_Toc88048910)

[3.5 Non-Functional Requirements 12](#_Toc88048911)

[3.5.1 Security and Privacy Considerations 12](#_Toc88048912)

[3.5.2 Business Continuity Requirements 12](#_Toc88048913)

[3.5.3 Business Volumes and Performance Expectations 12](#_Toc88048914)

[3.5.4 Archiving and data retention requirements 13](#_Toc88048915)

[4 Solution Design 13](#_Toc88048916)

[4.1 Network Infrastructure 13](#_Toc88048917)

[4.2 Key design decisions 15](#_Toc88048918)

[4.3 Components and Interfaces 16](#_Toc88048919)

[4.4 Security architecture 16](#_Toc88048920)

[4.4.1 Confidentiality 16](#_Toc88048921)

[4.4.1.1 Data in transit 16](#_Toc88048922)

[4.4.1.2 Data at rest 16](#_Toc88048923)

[4.4.2 Integrity 17](#_Toc88048924)

[4.4.3 Availability 17](#_Toc88048925)

[4.4.3.1 High availability 17](#_Toc88048926)

[4.4.3.2 Disaster recovery 18](#_Toc88048927)

[4.4.4 Backup plan 18](#_Toc88048928)

[4.4.5 Archiving and data retention 18](#_Toc88048929)

[4.4.6 Logging & Monitoring 18](#_Toc88048930)

[4.4.7 GDPR compliance 18](#_Toc88048931)

[4.4.7.1 Data Protection by Design and Default (DPDD) 18](#_Toc88048932)

[4.4.7.2 Targeted Bulk Erasure (TBE) Integration 18](#_Toc88048933)

[4.5 Operational architecture 19](#_Toc88048934)

[4.5.1 Operational responsibility 19](#_Toc88048935)

[4.5.2 Testing and Deployment 19](#_Toc88048936)

[4.5.3 Capacity planning 20](#_Toc88048937)

[4.5.3.1 ETL Specific capacity planning 20](#_Toc88048938)

[4.5.3.2 BAU Specific capacity planning 21](#_Toc88048939)

[4.5.3.3 Projected capacity planning 22](#_Toc88048940)

[4.6 Technical debt 22](#_Toc88048941)

# Solution Background

## Solution Context

The DSA department presently receives invoices through 2 routes:

* electronic JSON-based invoices via the e-Invoice platform, and
* PDF-based invoices via email.

Due to the volume of invoices received and the associated time to process them, there are two RPA automations running in production that execute daily. The Phase 2 automation deals with non-medical help (NMH) invoices which represent most invoices submitted. This automation presently has a success rate that varies between 30% to 40% depending on the quality of the underlying master data required for processing. This master data is presently held in over 100,000 user-specific spreadsheets and a handful of supplier spreadsheets.

There are several business drivers for improving the quality of this reference data, including:

* Increased automation success rate
* Reduced FTE costs for service provision
* SLA adherence
* Improved data governance

The solution being proposed is a database application to manage basic supplier data, student entitlement records and associated invoice submissions.

## Objectives

Create a brand-new database system to manage basic DSA supplier data, student DSA entitlements and associated invoice details against those entitlements.

The primary goal is to provide improved DSA invoicing data integrity to dramatically increase automation success rates, thereby reducing FTE costs associated with manual invoice processing.

Secondary goals of the project are:

* Improve assurance of customer data via the removal of files on shared directories
* Improve fraud prevention capabilities through a single searchable dataset
* Increase the scope of automation to include manual invoice submissions
* Dramatically reduce the risk of missed SLA for invoice payment

In terms of operational problems that are solved:

* Invoice processing still requires significant resource allocation. This is challenging operationally, particularly because the submission of invoices is highly volatile, with as few as 100 and as many as 6,000 being submitted on any given day.
* Extensive student entitlement and invoice data is held in unprotected files on shared network folders. Given that these entitlements and invoices are related to student disabilities, this is highly sensitive data to be held in such a fashion.
* Data gathering for fraud prevention activities can take weeks to consolidate because of the inconsistent file-based approaches to invoice submission and payment.
* Manual invoice submission is expensive and prevents unified reporting and fraud prevention activities as there is no invoice record in a digitally searchable format, unlike e-Invoice submissions.

## Assumptions

1. For the project to deliver against its cost-benefit analysis and produce a ROI in year 1, the project must utilise existing infrastructure where possible. To that end, it is assumed that the TSCI platform that was built to deliver services to Targeted Support, which includes the DSA function, will be the target platform for the management of the invoice data that supports automation.
2. It is assumed that the TSCI platform will be assured to the required security standards of the organisation before the solution moves into OAT/Production. The costs associated with this assurance and remediation work are not fully known at this juncture. However, from a budgetary perspective, we have 30 days set aside for such a review and 10 days set aside for previously identified encryption work.  
   **Update:** subsequent assurance work (DB Encryption, DAM) has been carried out successfully.
3. It is assumed that the Evolve initiative does not significantly alter its roadmap for Salesforce to include invoicing at an earlier juncture to the extent that it erodes the business benefit of the proposed invoicing solution.  
   **Update:** Replacement of E-Invoicing is not in scope for initial phases of the two DSA improvement initiatives underway presently.

## Risks

The following risks have been identified and mitigated.

Table 1 Risks

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **Description** | **Mitigation** | **Owner** |
| 1 | There is no dedicated allocation of security resources | We have accepted the risk and plan to engage with security to address this. **Update:** Security resource was obtained for the duration of the build. | Andy Neve |
| 2 | The long-term owner of the system within TG is unclear. | Whilst the long-term ownership is unclear, this will be mitigated by using commonly-used programming languages with abundant market resource availability (Java) coupled with an architectural pattern (Bob Martin’s Clean Code Architecture) that is widely used in industry and also within SLC (TBE architecture). | Andy Neve |

## Issues

Table 2 Issues

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Issue #** | **Description** | **Owner** | **Status** | **Comment** | **Section affected** |
| 1 | Platform encryption delay |  |  | The project has been delayed due to waiting for encryption on the TDCI platform, where this system is located, to be completed | QA and UAT testing of ETL data integrity. |

## Dependencies

The solution has dependencies as shown in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Number** | **Type** | **Title** | **Description** |
| 1 | Hard | Dependent on ETL process completion | Robots have been designed and built to transfer the tracker sheet data from the 100k+ tracker sheets that hold current customer information |
| 2 | Hard | Dependent on DSA Phase 2 automation refactoring | The current DSA automation which processes supplier invoices must be refactored to utilize the new system in order for a large portion of the project’s business benefit to be realised. |
| 3 | Hard | Dependent on successful upload of existing supplier information | A master spreadsheet has been compiled of all existing relevant suppliers which will be uploaded to the live system DB upon launch. |
| 4 | Soft | Dependent on clone of e-Invoicing database to TSCI platform | The business processes the system is designed to support, especially the refactored DSA automation, require that the planned clone of the e-Invoicing database onto the TSCI platform is completed successfully so that custom fields and API calls can be established. These will be necessary for an extended scope of the refactored DSA automation. |

## Out of Scope

|  |  |
| --- | --- |
| **Number** | **Functionality identified as out of scope** |
| 1 | In the DSA Apply-Assessment-Entitlement-Consume lifecycle, the Apply and Assessment stages are out with the scope of this project. Note however that modifications to affected automations are within the scope of the delivery. |
| 2 | Bulk transfer of existing Equipment and NAR entitlement and invoice data is out of scope of this project. In the ETL process, only entitlement and invoice data for NMH, General Allowance, and Travel type entitlements will be transferred over. This is due to the tracker sheets only holding information on these types of entitlements. |

## High-Level Delivery plan

Table 3 Delivery plan

|  |  |
| --- | --- |
| Indicative date | Milestone |
| Sep 2020 | Project Pre-discovery |
| Oct 2020 | Discovery |
| Nov 2020 | Design |
| Mar 2021 | Security Review |
| Apr 2021 | Build |
| Jun 2021 | Testing |
| Jul 2021 | Extract-Transform-Load Discovery |
| Aug 2021 | ETL Design |
| Sep 2021 | ETL Build |
| Oct 2021 | ETL Test  TSCI Platform Security Sign-off |
| Nov 2021 | TAB Review  ETL Deployment  Staff Training  Core Build Deployment |

# Requirements and Characteristics

## Business Processes Supported

The system will support recording, retrieving, and displaying customer entitlement, invoice, session, and Supplier information via a web-based UI. In the DSA Apply-Assessment-Entitlement-Consume lifecycle, this system will support business processes used to record and facilitate steps in the Entitlement and Consume stages. Current processes utilise Excel tracker sheets to record data, which will now be entered and edited in the new system.

Examples include:

* Creating and editing new customer entitlement records
* Entering and editing entitlements details against specific customers
* Entering and editing payment/invoice information against specific entitlements
* Entering and editing session information against specific payments
* Creating and editing Supplier records
* Tracking notes against and updating the status of supplier records

There is also a custom UI for entering invoice-level data which will then be sent to the e-Invoicing database which can then be passed along the payment business process that includes the present use of robotic process automation.

For an overview of key workflows, please see section 3.3.3.

## High-Level Functional Requirements

Requirement to allow business users to store information about suppliers:

* Supplier type, status, address, and relevant contact information
* Specific information depending on type of supplier

Requirement to allow business users to store customer information:

* Limited to customer CRN, domicile, course start, and end date, and entitlement start date

Requirement to allow business users to store entitlement information:

* Store entitlements against specific CRNs
* Record entitlements against specific suppliers when relevant to entitlement type
* Custom fields and dropdowns specific to type of Entitlement being entered
* Custom validation specific to type of entitlement
* Ability to quickly enter similar entitlement records or have records auto populated to fill out full course dates in line with common business processes

Requirement to allow business users to store track payments in an entitlement ledger

* Store line-item level payment entries against specific entitlements (which are tied to a specific CRN and supplier)
* Store invoice-level payment information against specific entitlement type (NMH)
  + Store line-item level information as sessions against these types of payment entries
* Custom fields and validation depending on entitlement type payment is created against
* Ability to record and update payment status, and associated changes in level of validation in line with status (i.e., Paid status requiring fullest validation steps)

Requirement for additional, custom UI for entering Invoices:

* Specific fields and validation depending on entitlement type invoice is intended for
* Invoice data to bypass DSA core database and be sent to e-Invoicing system database

### Business logic

High-level Business has been broken down in line with the requirements detailed above.

Supplier-relevant business logic:

* Duplicate supplier name checks prior to allowing the creation of new supplier records
* Updating supplier status in line with required minimum data depending on supplier type – i.e. HEP suppliers must have all general minimum fields provided plus HEP details in order to reach ACTIVE status.
* Locking supplier records with a status of ACTIVE\_AWAITING\_PAYMENT\_DETAILS after bank account details have been changed until an admin-level user approves the changes

Customer-relevant business logic:

* Duplicate CRN checks prior to allowing the creation of new customer records

Entitlement-relevant business logic:

* Duplicate checks when creating and updating NMH, non-reimbursement Equipment, Taxi, NAR, and General Allowance type entitlements – reimbursement Equipment, mileage, and taxi reimbursement type entitlements do not have duplicate checks to mirror necessary business processes.
* Date range checks when creating and updating entitlements to ensure entitlements cannot be valid after course end date
* Soft stops (UI pop up which can be overridden by user with audit trail) when:
  + Entitlement date ranges pre-date customer entitlement start date
  + Cost entered is out with recorded cost range for specified NMH type for NMH entitlements
  + Cost entered differs from cost recorded on NAR supplier record when creating/editing NAR entitlements

Payment-relevant business logic:

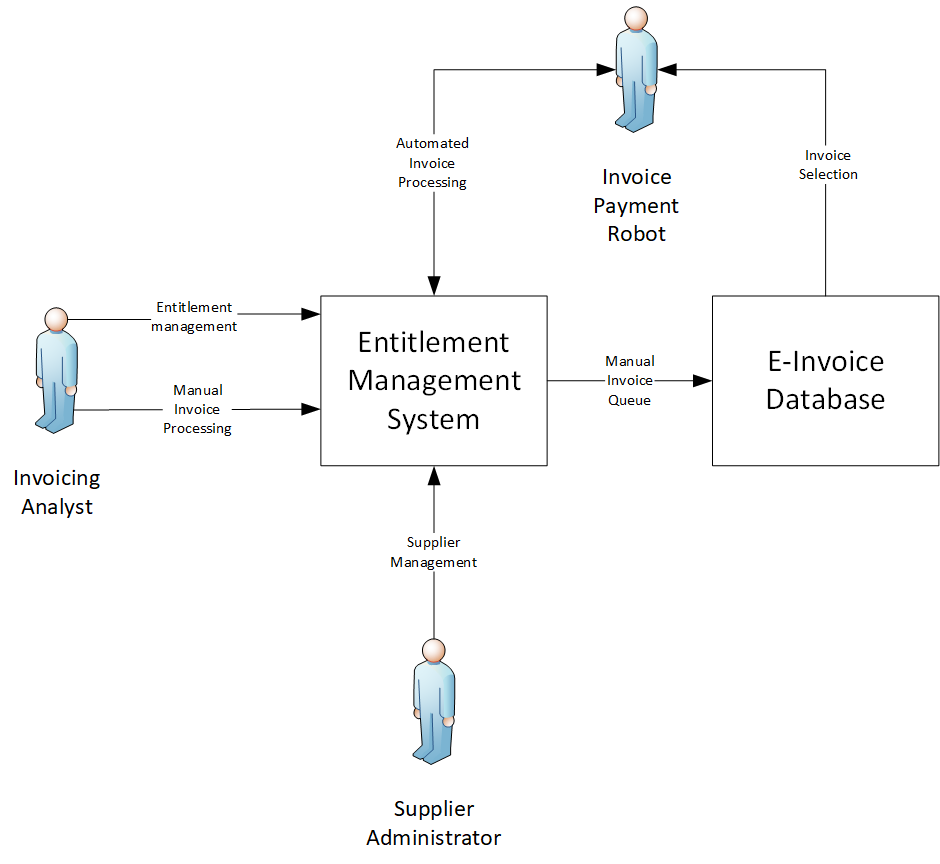
* Entitlement breach checks when a payment is created/edited with PAID status
* Payment duplicate checks across NMH, non-reimbursement Equipment, Taxi, NAR, type entitlements
* Session duplicate checks for ATTENDED status sessions against NMH type payments

### Integration

1. The DCI SFD service pulls real-time information (Customer First/Last name) for read only purpose from the Customer web service on the Assess platform.

## Information Architecture

### Information Context

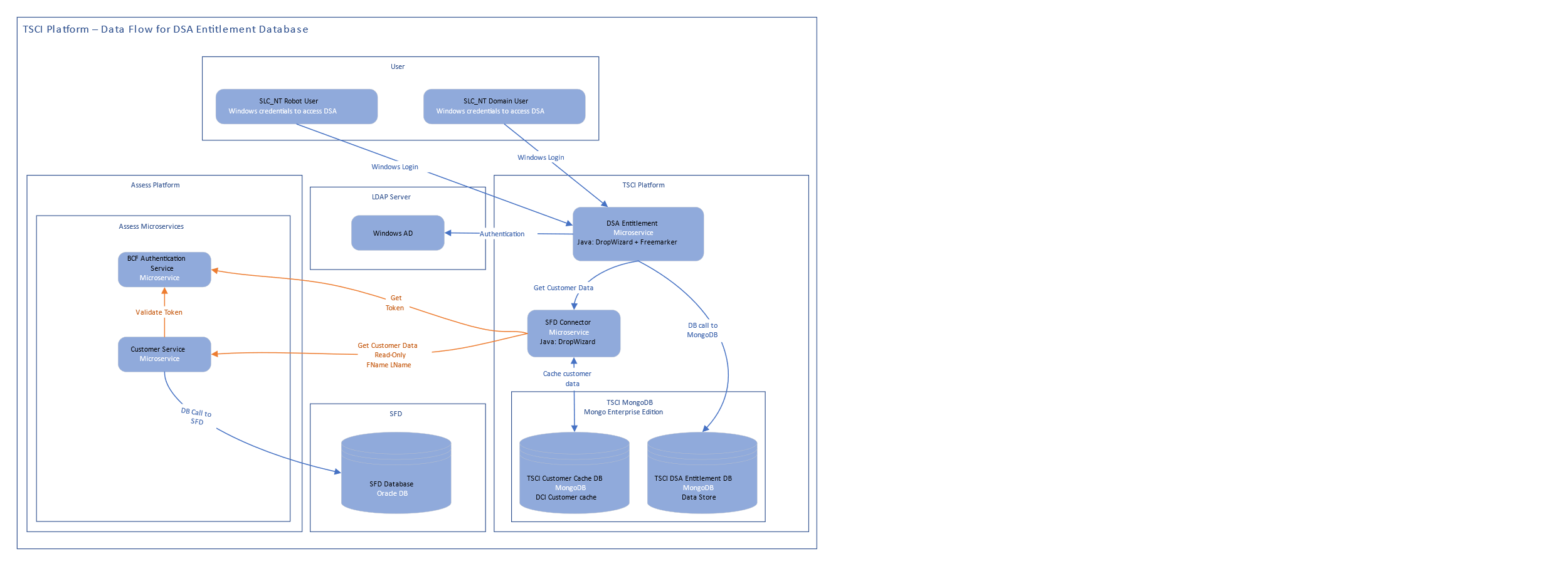


### Information Entities

Table

Description automatically generated with medium confidence

### High Level Workflows





## User community description

|  |  |  |
| --- | --- | --- |
| **User Group** | **Usage Reason** | **Comment** |
| DSA Invoicing Analyst | Carrying out the Entitlement and Consume stages of the DSA Apply-Assessment-Entitlement-Consume lifecycle | Utilising system’s dedicated UI |
| DSA Supplier Management | Maintaining the supplier database within the entitlement system, | Utilising system’s dedicated UI and reporting functionality |
| Robotic Process Automation | Carrying out Payment Ledger activities for invoice processing | Calling dedicated API endpoints with robot credentials |

## Non-Functional Requirements

Non-functional requirement information can be found in from P000468 DSA - Security and Governance NFRs v0.7 document provided by Security PRR.

### Security and Privacy Considerations

The (potential) PII held within the system is as follows:

* CRN
* Entitlement Type (the customers broad category of disability could be inferred from this entity)

### Business Continuity Requirements

* Service hours – 07:00 to 02:00 Monday to Friday (19 hours per day, 399 hours per month)
* Scheduled Maintenance notice period > 24 hours
* Service Hours Availability > 99.5% (equivalent to < 2 hours downtime per month)
* MTBI > 30 days (i.e. less than one P1/P2 service interruption per month)
* Disaster RTO <= 1 business day
* Disaster RPO <= 1 business day

### Business Volumes and Performance Expectations

Section 4.6.3.2 lists the full capacity requirements for BAU activity on an annual basis and is relatively accurate given that typical invoice numbers and entitlements are known.

There is no anticipated growth beyond existing usage patterns based on log data retrieved from robotic processing over the last 2 years. Similarly, at the time of writing, there was no known policy change likely to impact these trends.

Regarding peak processing, invoice submission is relatively stable from September through June, whereas July and August are relatively quiet as disability students have reduced need for support outside of term times.

From a transaction distribution perspective, invoices are processed by robots on weekdays, typically from later afternoon through to the evening, although this changes based on robot capacity and availability that is impacted by other more time-critical automations.

Transaction volumes are inconsistent based on suppliers’ preferences for submission time and frequency of invoice submission through the e-Invoice portal and via post for manual processing.

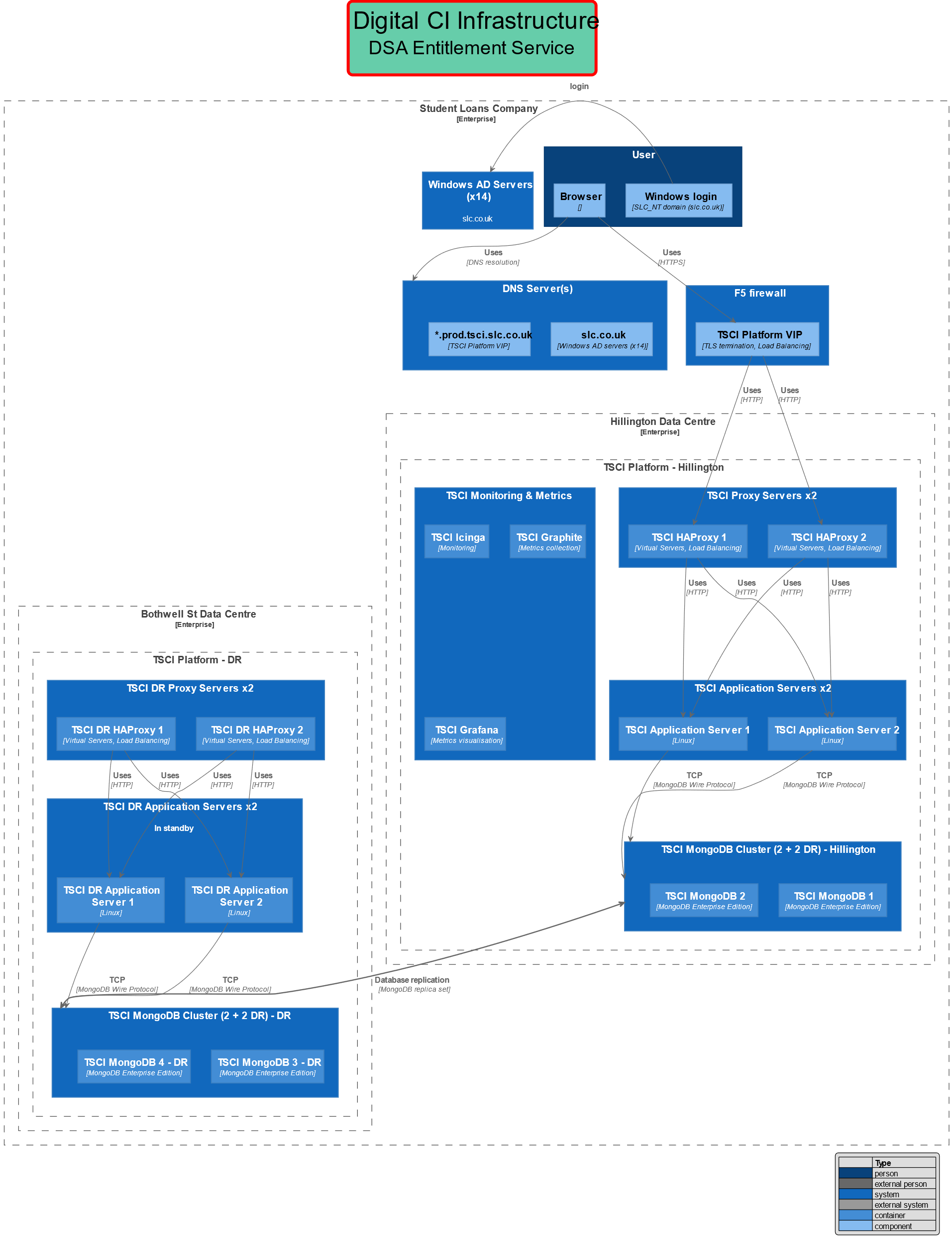
Reporting requirements have yet to be formally specified and will be defined as part of a subsequent delivery. However, one such potential use case for large scale reporting is the capability to analyse stated support consumption for potential fraud activities.

### Archiving and data retention requirements

Retention periods for DSA entitlement data will be maintained as per the guidelines for data retention in the DSA space in line with policy therein. It is not anticipated that any additional archiving needs exist. As such, the data will be retained on the TSCI platform across the 4 replica set Mongo database configuration, until such a time that the system is integrated into the TBE programme. This has already been discussed with said programme and is scheduled for integration at a time that suits the programme.

# Solution Design

## Network Infrastructure



## Key design decisions

| Subject Area | Security |
| --- | --- |
| Design Decision | User Authentication |
| Issue or problem statement | A secure mechanism for user authentication is required that does not introduce complexity or is cumbersome to register new users. |
| Assumptions | Users will already have a Windows AD account (SLC\_NT domain). |
| Alternatives | 1. BCF Authentication 2. Windows AD |
| Decision | 2. Windows AD. |
| Justification | Simplifies user administration. Simplifies integration. |
| Implications | User groups required in AD. |

| Subject Area | Customer data access |
| --- | --- |
| Design Decision | Use existing TSCI SFD Connector component that links to Assess CustomerService. |
| Issue or problem statement | The solution requires read-only access to some customer data from the SFD database. |
| Assumptions | None. |
| Alternatives | 1. Use Assess CustomerService to retrieve data. 2. Use existing TSCI SFD Connector microservice to retrieve and cache data from SFD database    1. access via Assess CustomerService (currently working)    2. access SFD directly (requires security approval + additional testing) 3. Create direct connection to SFD database |
| Decision | Option 2/i. |
| Justification | Reuse of a proven solution already used by Migrant Workers DB and Estrangement DB. |
| Implications | Additional dependency: TSCI SFD Connector. |

## Components and Interfaces

* TSCI / DCI [SFD Connector](http://gitlab.slc.co.uk/ddts/dci-sfd)
* Existing SLC Microsoft Active Directory implementation - domain: SLC\_NT. (Highly available.)

| **Type** | **Name** | **Description** | **Status** | **Owner** |
| --- | --- | --- | --- | --- |
| REST API | DCI SFD connector service | Retrieve customer data from SFD database via Assess Customer Service + cache these data  Swagger spec – changed for Allow new service to connect | Changed | Operation CI |

## Security architecture

### Confidentiality

Database encryption has been implemented across all open environments – OAT, PROD.

Granular AD groups have been implemented and full front and back-end functionality-specific verification has been implemented in line with user roles.

All data are stored securely in password protected MongoDB databases. Direct access to the database servers is only possible from Assure Linux bastion host. Access to the database via Mongo OpsManager is password protected. In the application logs, only identifiers are logged (such as CRNs) – logging of details is switched off by default (although it is possible to change the logging level for debugging purpose temporarily, if it is unavoidable to track errors).

#### Data in transit

Data in transit is encrypted via SSL

#### Data at rest

Encrypted using Vormetric encryption software

### Integrity

Any access to data via the application (either to read or to modify) is saved in the audit log. The database is password protected; direct database access to the data is restricted: only possible from Assure Linux bastion host or via Mongo OpsManager.

### Availability

The system will reside on the TSCI platform. This platform has been running for a number of years and has load-balanced application and database servers with multi-site real-time replication.

#### High availability

The application if highly available:

* F5 load balancer
* Two haproxy nodes to provide routing and load balancing to TSCI services
* Two application server nodes to host the TSCI services
* Four node MongoDB cluster stretched between Hillington and Bothwell (for DR)
* Highly available Assess Platform CustomerService
* Existing 14 node Windows AD cluster (for authentication / authorisation)

Service monitoring and metrics collection is provided by the TSCI MM infrastructure:

* [TSCI Monitoring - Icinga2](https://monitoring.prod.tsci.slc.co.uk/)
* [TSCI Metrics - Grafana](https://metrics.prod.tsci.slc.co.uk/)

For DR, the following server nodes are in standby (in Bothwell St):

* Two DR haproxy nodes to provide routing and load balancing to TSCI services;
* Two DR application server nodes to host the TSCI services;
* Two DR nodes of the four-node MongoDB cluster.

All above nodes are kept in sync with production:

* DR MongoDB nodes are part of the replica set;
* Configuration of DR haproxy nodes are kept up to date, in sync with the production haproxy nodes;
* All services that are deployed to the production environment are also deployed to DR application servers, with the correct configuration.

DR procedures (manual steps required):

* Reconfigure MongoDB cluster to work with two servers;
* Start the services on the two DR application servers;
* Redirect network traffic in DNS servers to the DR environment.

#### Disaster recovery

TSCI platform has a multi-site 4 database replica set for disaster recovery and BC.

### Backup plan

Regular MongoDB backups are configured in Mongo OpsManager:

* one snapshot daily is retained longer
* 6-hourly snapshots (retained for a week)

### Archiving and data retention

No archival procedure has been established.

Data retention:

* no automated data deletion has been implemented;
* users with admin credentials can remove individual applications using the UI;
* using the API, deletion can be programmatically triggered if necessary.

### Logging & Monitoring

Service monitoring and metrics collection is provided by the TSCI MM infrastructure:

* [TSCI Monitoring - Icinga2](https://monitoring.prod.tsci.slc.co.uk/)
* [TSCI Metrics - Graphite + Grafana](https://metrics.prod.tsci.slc.co.uk/)

Logging is local only, with daily log rotation. Default log level is INFO - limited information is logged if there are no errors (only IDs like CRN are recorded). In case of errors, enough information is logged to reproduce the issue. Logs are retained forever unless manually removed.

### GDPR compliance

* Subject access requests can be accommodated by using the Reporting functionality (by admin users);
* Incorrect data can be corrected by assessor or admin users (in an audited way);
* Information can be completely removed by users having admin role

#### Data Protection by Design and Default (DPDD)

DPDD was enacted with guidance from Security personnel as part of the system’s security review.

#### Targeted Bulk Erasure (TBE) Integration

Should be no impact on TBE at this instance as, prior to creation in DSA Core database, all customer records will have had an initial needs assessment payment processed against their CRN. Therefore, data is valid to store without considering TBE architecture.

Initial engagement with the TBE project has taken place and DSA entitlement data is on TBE backlog for future integration.

## Operational architecture

### Operational responsibility

Table 7 Operational Responsibility

| Component | Ownership | Operational responsibility | Comment |
| --- | --- | --- | --- |
| Windows AD | Current operations team | Monitoring, 1st level support | Re-used component |
| F5 nodes | Current networks team | Monitoring, 1st level support | Re-used component |
| Assess customer service | Assess | Monitoring, 1-3 level support | Re-used component |
| Application servers, MongoDB, Monitoring | NOC Team | Monitoring, 1-3 level support | components |
| Application servers, MongoDB, Monitoring | Operation CI | 2-3 level support | components |

### Testing and Deployment

**Deployment and test strategy**

1. Automated build and deploy GoCD pipelines have been created, along with deployment scripts, and Ansible codes for configuration. Ansible codes are in separate projects within the same GitLab repository subgroup. Deployment scripts are stored within the respective GitLab project: /go-artifacts/run\_pipe.sh
2. On source code changes, the **build pipeline** executes the unit tests and integration tests.
3. On success, the service is packaged as RPM and published on dc1repo.slc.co.uk.
4. The RPM package is deployed by the GoCD pipeline to **DEV environment**; the configuration files are deployed by Ansible code.
5. The service is automatically started by the deployment code and verified by the application's health check.
6. The deployment is further validated by automatically executing the API automation tests that exercise all API functions in a reproducible way.
7. Validating the UI functionality is semi-automatic: Selenium based Katalon test scripts have been created and stored in a[GitLab repository](https://gitlab.slc.co.uk/ddts/dsa-core/dsa-core). These test scripts exercise the full range of UI functionality and validate that the experienced behaviour corresponds to the expected. The execution of these test scripts is triggered manually by the developers.
8. On successful deployment to DEV environment, the deployment will automatically continue to the **NextTest environment**, with similar steps as for DEV - but without Selenium/Katalon UI testing.
9. Deployment to **OAT environment** is triggered manually. The deployment starts on one OAT node, then a health check and API automation testing is performed on that node. On success, the deployment continues similarly on the other node. This way, the service is kept available during deployment via load balancing: this is an "uptime deployment".
10. Deployment to **PRODUCTION environment** is triggered manually. The deployment affects one node at a time, followed by a health check (no automation testing is executed in PROD); on success, the deployment continues on the other node. This way, the service is kept available during deployment via load balancing: this is an "uptime deployment".
11. On successful deployment to PROD, the same version is automatically deployed to ***DR environment*** (but the service is not started there).

Graphical user interface, application

Description automatically generated

**Testing Plans:**

Unit tests, integration tests, API and UI automation tests have been implemented. For more details, see above in "Deployment and test strategy" section.

Furthermore, extensive manual testing has been performed, both by the development team and the UAT team

### Capacity planning

#### ETL Specific capacity planning

Storage needed for initial ETL process is based on estimates made from total customer base to be migrated.

Based on an initial 100k customers identified for ETL, it is assumed that 60% will pass ETL filtering to determine if recent (19/20 onwards) academic years are included in customer tracker sheet. Sessions are not in scope for ETL transfer

Total storage estimates are based on individual object size only.

ETL Storage volumes

**Customer:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Supplier | 2350 | 2824700 | bytes |
|  |  | 2.69384 | MB |
|  |  | 0.00263 | GB |

**Customer:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Customer | 60,000 | 15660000 | bytes |
|  |  | 14.9345 | MB |
|  |  | 0.0146 | GB |

**Entitlement:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Entitlement | 120,000 | 44040000 | bytes |
|  |  | 41.9998 | MB |
|  |  | 0.041 | GB |

**EntitlementPayment:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| EntitlementPayment | 480,000 | 214560000 | bytes |
|  |  | 204.62 | MB |
|  |  | 0.1998 | GB |

**Total**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Supplier | 2350 | 0.00263 | GB |
| Customer | 60,000 | 0.01458 | GB |
| Entitlement | 120,000 | 0.041 | GB |
| EntitlementPayment | 480,000 | 0.1998 | GB |
|  |  | 0.258 | Total GB |

#### BAU Specific capacity planning

Annual estimates are provided below of ongoing storage requirements from BAU activities.

Assumptions:

Suppliers:

* Amount onboarded annually is negligible

Estimates for entitlementPayment total is made on following assumptions:

* 500k total invoices per annum
* 80% are NMH invoices which required only one entitlementPayment entry per invoice
* 20% are non-NMH which required at least 1 entitlementPayment entry per invoice, for an average of 2

Session estimates are based on following assumption:

* 80% of all invoices (500k annual invoices) are NMH type
* Approx 3 sessions recorded per NMH invoice

**Customer:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Customer | 50,000 | 13050000 | bytes |
|  |  | 12.44544983 | MB |
|  |  | 0.012 | GB |

**Entitlement:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Entitlement | 100,000 | 36700000 | bytes |
|  |  | 34.99984741 | MB |
|  |  | 0.034 | GB |

**EntitlementPayment:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| EntitlementPayment | 600,000 | 268200000 | bytes |
|  |  | 255.7754517 | MB |
|  |  | 0.25 | GB |

**Session:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Session | 1,200,000 | 346800000 | bytes |
|  |  | 330.734 | MB |
|  |  | 0.32 | GB |

**Total**

|  |  |  |  |
| --- | --- | --- | --- |
| **Record type** | **Volume** | **Total** |  |
| Customer | 50,000 | 0.012 | GB |
| Entitlement | 10,000 | 0.034 | GB |
| EntitlementPayment | 500,000 | 0.25 | GB |
| Session | 1,200,000 | 0.32 | GB |
|  |  | 0.616 | Total GB |

#### Projected capacity planning

Adjusted projections assuming a 10% annual growth in overall volume based on historical data

**Total**

|  |  |
| --- | --- |
| **Academic year** | Total Annual GB Useage |
| 2021/22 | 0.616 |
| 2022/23 | 0.6776 |
| 2023/24 | 0.745 |
| 2024/25 | 0.82 |
| 2025/26 | 0.902 |

**Total data stored including initial ETL**

|  |  |
| --- | --- |
| **Academic year** | Total GB stored |
| 2021/22 | 0.874 |
| 2022/23 | 1.5516 |
| 2023/24 | 2.2966 |
| 2024/25 | 3.117 |
| 2025/26 | 4.02 |

## Technical debt

Managed to avoid technical debt through ensuring that codebase is configurable to allow for maximum extensibility and maintainability for anticipated CI activities including:

* Updating changing VAT rates
* Updating changing Support Types and associated support type rates
* Adding and changing class fields/names