

DATALYNX DATA MANAGEMENT SUITE

Technical Overview



September 2014



Contents

1. INTRODUCTION.....	3
1.1. ENTERPRISE DATA MANAGEMENT	3
1.2. PRODUCT OVERVIEW	4
1.3. BUSINESS SOLUTION OVERVIEW	6
2. DMS SYSTEM ARCHITECTURE	8
2.1. TOPOLOGY	9
3. KEY FEATURES	10
3.1. BROAD CONNECTIVITY & INTEGRATION.....	10
3.2. INBUILT PROCESSING FUNCTIONS	10
3.3. CODE-FREE CONFIGURATION	11
3.4. REUSABILITY	11
3.5. SCALABILITY & PERFORMANCE	12
3.6. SECURITY & ADMINISTRATION	13

1. INTRODUCTION

The Datalynx Data Management Suite (DMS) is a next-generation, fully integrated data management software solution that helps to maximise the business value organisations derive from data and information assets.

Datalynx DMS is used by clients, integration partners and the Datalynx Professional Services Group as the technology that underpins enterprise data governance initiatives. Datalynx invests in the ongoing development of the Suite to support evolving data management requirements and address the challenges faced by organisations.

1.1. ENTERPRISE DATA MANAGEMENT

Integrated Technology

The Datalynx DMS is a fully integrated enterprise data management solution that delivers all the capabilities needed for effective management and utilisation of corporate data assets. Supporting both project-based and operational (business-as-usual) activities, the DMS enables the development and deployment of comprehensive enterprise solutions via a single product suite.

Furthermore, the DMS enables clients to overcome the common challenges of trying to integrate multiple disparate technologies to achieve business outcomes.

Business Solutions

Designed to help organisations resolve real world business problems and realise the benefits of effective data management, the DMS supports a comprehensive set of business capabilities, including:

- ✓ Data Quality Management
- ✓ Business Intelligence / Data Warehousing
- ✓ Data Migration
- ✓ Master Data Management
- ✓ Metadata Management
- ✓ Data Auditing & Fraud Detection
- ✓ Enterprise Data Integration
- ✓ Data Archiving
- ✓ Data Management for the Cloud
- ✓ Big Data Integration

1.2. PRODUCT OVERVIEW

The Datalynx DMS comprises two integrated software systems:

- a. Datalynx Data Xplorer
- b. Datalynx Data Xchange

These applications work together to address the most sophisticated business requirements by delivering solutions comprising a combination of core data management capabilities.

Datalynx Data Xplorer: - comprehensive data analysis, validation and reporting.

Capability	Description
Data Profiling (Analysis)	Analysis of data structures, characteristics and values. Identification and reporting of broad range of data quality issues and anomalies, including inconsistencies, missing values and duplicate records.
Data Validation	Column-based evaluation of data against business rules to identify exceptions and non-compliant transactions. Utilises deterministic and probabilistic (fuzzy) matching algorithms as required.
Monitoring	Automated evaluation of profiling criteria and business rules on a scheduled basis to identify issues and transactions of business interest.
Reporting	Production of profiling, analysis, verification and validation reports.

Data sets can be “registered” with Data Xplorer to expose their metadata and data values for analysis. Data Xplorer is designed for columnar analysis and includes a range of pre-built, reusable functions that can be applied to perform common profiling tasks such as:

- Duplicate checking
- Integrity verification
- Statistical analysis
- Data quality checking

It is possible to extend Data Xplorer’s inbuilt functions with reusable custom functions for performing business-specific analysis.

All functions are created using standard SQL and a group of related functions can be saved and processed as a “job” that can be rerun as required.

In addition to persisting the results of analysis for future comparison, Data Xplorer provides comprehensive reporting capabilities and the option to export data to Microsoft Excel and Microsoft Word.

Datalynx Data Xchange: - agile data management, ETL / ELT, data quality, migration & integration.

Capability	Description
Data Extraction	DMS integrates with any structured data source (RDBMS, files) using in-built connectors and supports the extraction of complete data sets from one or multiple sources, as well as ongoing changed data capture for synchronising data between systems.
Data Validation	Comprehensive row-based evaluation of data against business rules. The DMS rules engine can evaluate any number of business rules / criteria for a single data attribute, as well as perform analysis across multiple attributes, tables and datasets. A combination of deterministic and probabilistic (fuzzy) matching algorithms can be applied as required.
Data Cleansing / Quality Enhancement	Utilising the extensive set of inbuilt data quality / data manipulation functions provided with the DMS, it is possible to apply sophisticated business rules to cleanse, standardise and enrich data for use within an existing system, or for migration to a new application. The integrated DMS scripting engine also allows the addition of custom logic if required.
Data Transformation	The rich DMS data transformation functionality enables data structures in the source data sets to be prepared and aligned with those expected by the new (target) systems. Includes the creation of “virtual” fields, derived / calculated values, merged attributes and default values.
Data Migration	Data can be migrated “as-is”, or in combination with data analysis (profiling), quality enhancement and transformation processing. DMS supports migration between different platforms, database types & data structures and can also use delimited files as a data source as required.
Data Integration	The Datalynx DMS can be configured to serve as a sophisticated enterprise Data Processing Hub that provides integration between different enterprise systems, integration with BI / EDW and Big Data solutions, as well as integration between external and internal data sets.
Audit Tracking	The DMS maintains a comprehensive audit trail of the processing it performs to support transparency, verification and reporting requirements. With inbuilt version control it is possible to know at any stage how data has been processed from its original state right through to its final format and content.
Reporting	Reporting for data analysis and process verification is a key capability of the DMS. A range of comprehensive reports are provided “out of the box” and can be customised as required.

1.3. BUSINESS SOLUTION OVERVIEW

Discrete Solutions:

The following is a list of discrete business solutions that are available via the Datalynx DMS:

- i. Data Analysis (profiling):**
 - Issue identification
 - Data auditing / Fraud detection
 - Validation / verification & Reporting

- ii. Data Quality Management**
 - Project based data cleansing
 - Ongoing automated data quality monitoring and reporting
 - Data quality enhancement and standardisation for BI and Big Data solutions

- iii. Data Migration**
 - Legacy data migration
 - Data consolidation
 - Data Archiving

- iv. Master Data Management**
 - Establishment and ongoing management of single sources of truth for any data domain, including customers, products, vendors, locations etc.

- v. Metadata Management**
 - Comprehensive metadata capture, storage and reporting
 - Data lineage

- vi. ETL / ELT**
 - Incorporates *extract-transform-load* and *extract-load-transform* processing models.
 - Out of the box integration with broad range of data sources (See Section 3.1 for a list)

- vii. BI / Data Warehouse Solutions**
 - Highly scalable, end to end BI solutions incorporating data extraction, validation, quality enhancement / standardisation, MDM, data mart / EDW population & maintenance for integration of analytics/reporting tools.

- viii. Cloud Data Management**
 - Transition to standard platforms in the cloud
 - Data rationalisation
 - Data quality management
 - Ongoing monitoring and reporting

- ix. Big Data Integration Solutions**
 - Data extraction, preparation and staging for integration with Big Data technologies.

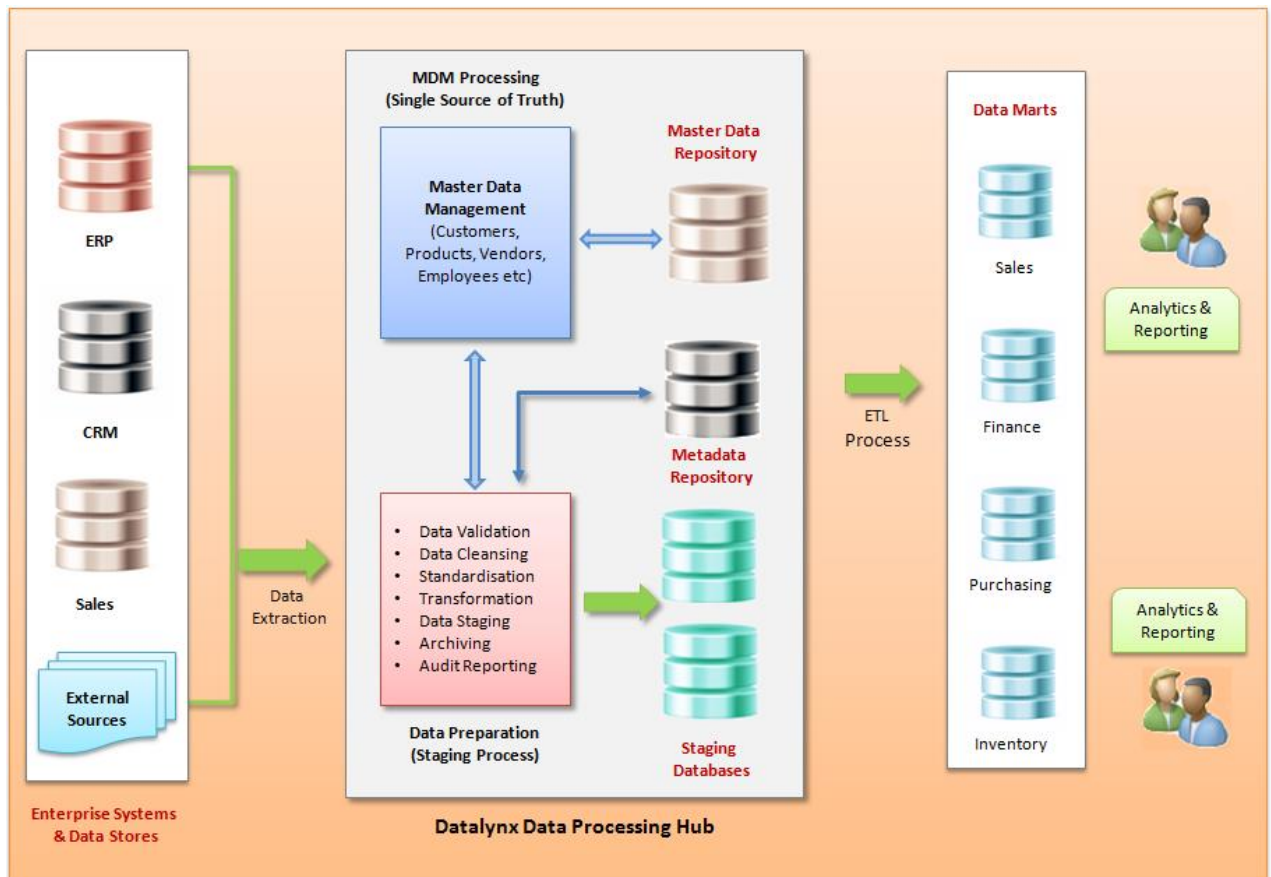
Datalynx Data Processing Hub Solutions

A key strength of the Datalynx DMS is its ability to combine any number of the individual data management capabilities listed in section above into an enterprise data management solution.

An integrated Datalynx Data Processing Hub can incorporate any combination of:

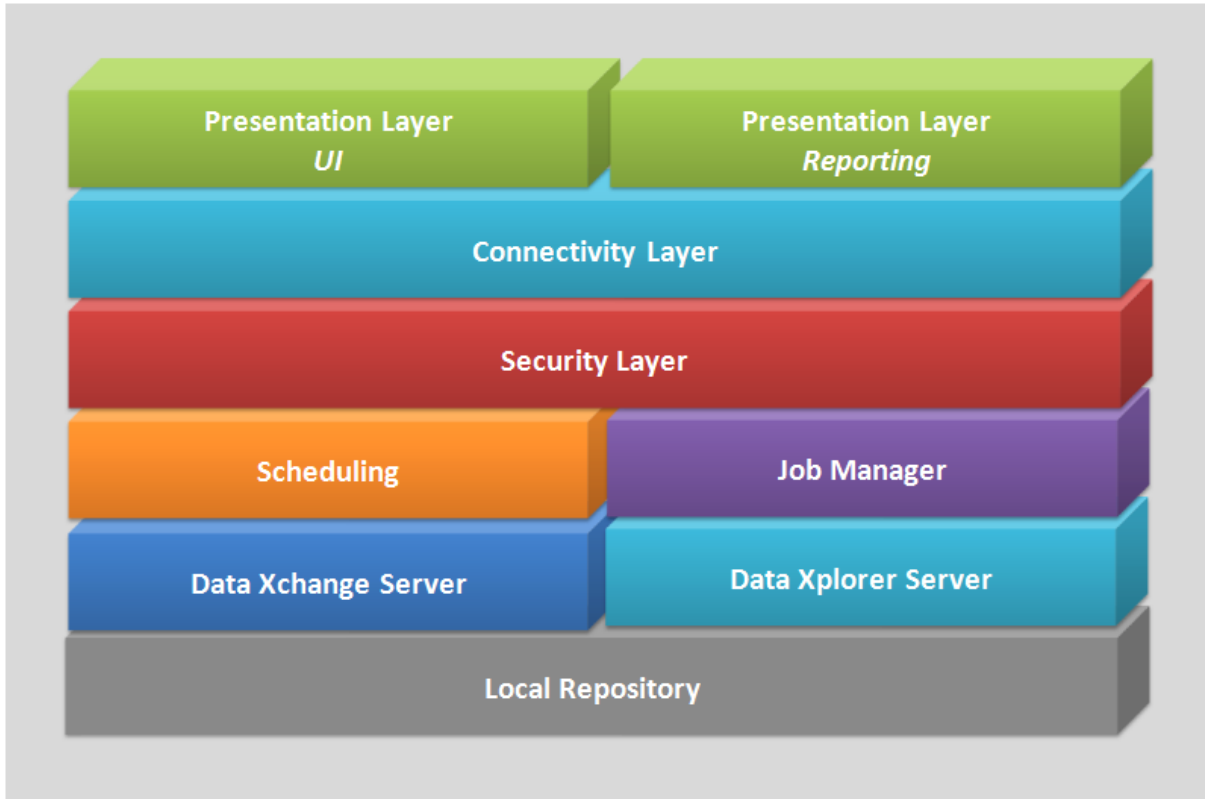
- ETL / ELT processing
- Data Quality Management
- Master Data Management
- Metadata Management
- Data Migration
- BI / EDW integration
- Big Data Integration
- Data Archiving
- Auditing and Fraud Detection

The diagram below demonstrates how the Datalynx Data Processing Hub can be configured to provide a comprehensive BI solution:



2. DMS SYSTEM ARCHITECTURE

The Datalynx DMS is comprised of the following key building blocks:



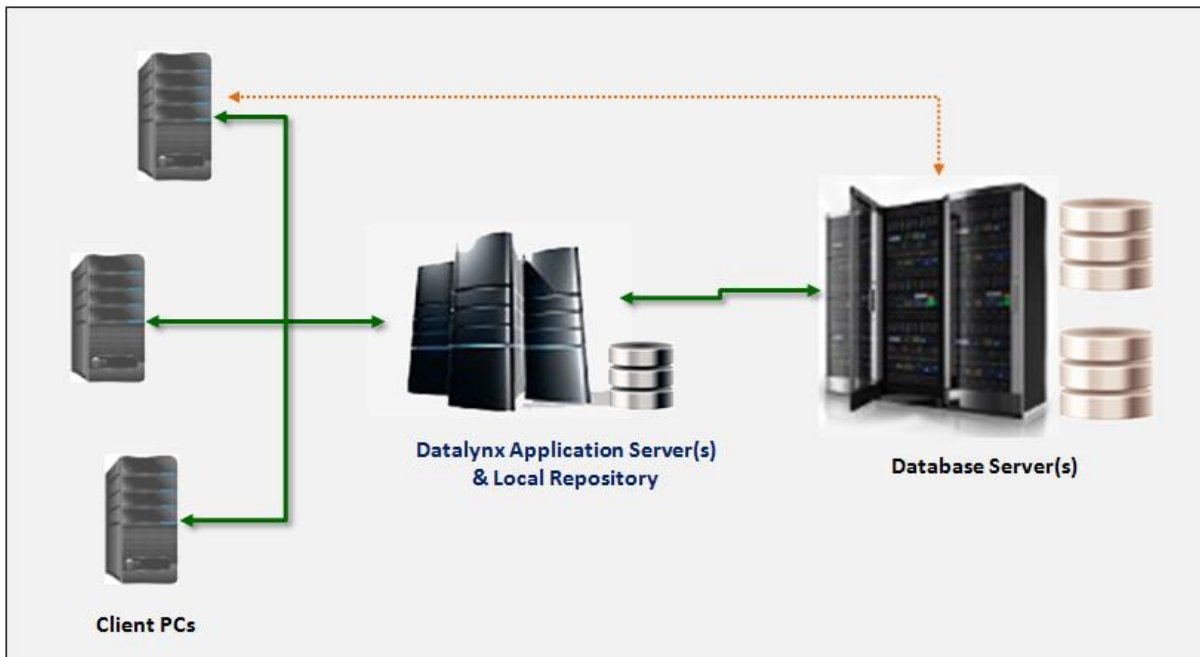
Component	Description
Presentation Layer - UI	The graphical user interface is presented as a Windows client application. Multiple sessions can be run concurrently on a single client.
Presentation Layer - Reporting	DMS reporting is delivered via a combination of Microsoft Word and Microsoft Excel report templates. Data extracts are provided as .csv files.
Connectivity Layer	The Connectivity Layer utilises native drivers, as well as OLE DB and ODBC to connect to a range of databases and structured file types.
Security Layer	The DMS has integrated security capabilities providing flexible access control and privilege management functions.
Scheduling	The job scheduling function allows for default immediate job processing as well as a visual calendar for scheduling when jobs will run, as well as their frequency.
Job Manager	The inbuilt Job Manager performs job prioritisation, server allocation and job queue management functions.

Component	Description
Data Xchange Server	The application server for the Data Xchange system.
Data Xplorer Server	The application server for the Data Xplorer system.
Local Repository	The local application database used for the operations of the DMS , as well as persisting results of data analysis, data processing and the audit log.

2.1. TOPOLOGY

The diagram below displays the standard DMS topology comprising desktop PCs for the DMS clients, Application Servers for the DMS Servers and Database Servers for the source / target business data sets.

- The desktop PCs run the DMS client (Data Xplorer / Data Xchange) and provide the windows based user interface.
- The Datalynx Application Server(s) are Windows physical or virtual servers that run the DMS Server software and house the local database repository.
- The Database Servers can be on any platform and are accessed by the Client PCs and the Datalynx Application Servers.



3. KEY FEATURES

3.1. BROAD CONNECTIVITY & INTEGRATION

The Datalynx DMS is preconfigured to connect with and between numerous databases and structured data sources and can be easily extended to add new sources as required.

DMS standard data sources comprise:

Oracle RDBMS	Microsoft SQL Server	IBM DB2
SAP HANA	Teradata	Sybase
MS Access	UniVerse	MS SQL Azure
Informix	MySQL	MS Excel
XML Files	Delimited Text Files (e.g. .csv)	.DBF, Paradox, Foxpro

The DMS also connects with any data source that has an *OLE DB* or *ODBC* connector and can integrate with proprietary data sources using the DMS custom Plug-in interface.

3.2. INBUILT PROCESSING FUNCTIONS

The Datalynx DMS includes an inbuilt library of over 100 optimised data processing functions that significantly enhance productivity and eliminate the need for custom programming. By being tightly integrated with the DMS rules engine, the functions deliver outstanding performance for complex processing tasks, including de-duplication, integrity checking, reconciliation etc.

Function type categories comprise:

- String functions
- Date & Time
- Logical & Binary
- Mathematical
- Regular Expressions (including address parsing)
- Miscellaneous (including fuzzy matching)

These functions can typically address very sophisticated data processing requirements. The functionality of the inbuilt functions is augmented by the DMS Scripting Engine that supports complex logic.

Test While Configuring

Another risk management and productivity enhancing feature is the ability to test the configuration as it is being built. The DMS provides access to a copy of the underlying data that can be used to test and verify the results of any function as it is being incorporated, thereby fast-tracking the design process and reducing the risk of unanticipated outcomes.

3.3. CODE-FREE CONFIGURATION

The Datalynx DMS does not require custom programming or coding in any proprietary languages.

The DMS enables designers to adopt an agile approach to creating business solutions by supporting an iterative development process for configuring jobs and tasks. This capability allows solutions to evolve to accommodate changes in requirements and address broader business needs over time.

The key business benefits of the configuration-based approach utilised with the DMS, include:

- High-productivity
- Reduced risks
- Significant reusability
- Reduced timeframes and lower costs for delivering solutions

The DMS has been designed to be used by ‘technical’ business analysts, in addition to more technically-focused data analysts and developers. It is not essential for a user to have a software development background, as long as they have a good understanding of database / data management concepts and mid-level SQL skills.

3.4. REUSABILITY

Reusable Functions

In addition to a large number of “out of the box” functions to perform key data management tasks, Data Xplorer can be extended with custom SQL functions for specific processing. Examples of inbuilt functions include:

- Date range and validity checking
- Duplicate record checks
- Data value frequency counts
- Minimum and maximum values
- Pattern validation (e.g. ABN, telephone number checking)

All functions can be parameterised, allowing them to be applied as many times and to as many data attributes as required.

Furthermore, once a function or set of functions has been developed, it is available to be reapplied in future. For example, the data quality profiling functions developed for a migration from Lotus Notes to Sharepoint can all be reused the next time a similar migration is to be undertaken. This capability has a significant impact on projects by reducing development risks & duration, as well as project costs.

Note: functions can be processed individually, or in saved ‘jobs’ comprising a set of related functions.

Data Xchange Templates

The Data Xchange system uses a wizard-like interface for building job processing configuration specifications called “templates”. This approach provides an intuitive framework within which very sophisticated data processing tasks can be defined.

Templates consist of four configuration tabs, used for defining:

1. Job description, security and execution parameters
2. Data source(s) definition and connection parameters
3. Data target(s) definition and connection parameters
4. Job configuration details, including:
 - a. Source to target field mapping
 - b. Validation rules and exception processing
 - c. The processing sequence within the template, as well as calls to linked templates
 - d. Data manipulation functions (cleansing, transformation, calculations, data standardisation etc.)
 - e. Job scheduling
 - f. Job management (priority, processing server allocation, tracking)

Data Xchange templates are version-controlled and logged against jobs as they are processed. This provides a clear audit trail of the type of processing applied to data to achieve particular outcomes.

Templates are persistent and so can be scheduled to run at a required frequency and at specified times. Templates can also be exported and shared with other users, supporting an effective model of centralised development and decentralised processing.

3.5. SCALABILITY & PERFORMANCE

The Datalynx DMS has been designed as an enterprise data management system and incorporates the necessary features to provide the highest level of performance.

Key features for scalability and performance include:

- a. In-memory processing model for enhanced throughput.
- b. The ability to address specific CPUs on a server, as well as any number of CPUs (up to a total of 32 CPUs on a single server).
- c. Inbuilt job management for automatically assigning complex or high-priority jobs to the appropriate server for processing.
- d. Concurrent processing of a single job across multiple servers, reducing the processing time for large datasets.
- e. Multithreading for increased and more efficient processing performance.
- f. Multiple concurrent DMS client sessions to enable several discrete tasks to be performed at the same time.

3.6. SECURITY & ADMINISTRATION

The DMS employs a powerful security framework for managing user access rights and privileges, incorporating multiple security levels and caveats.

Security is applied to:

- System Logon / Access
 - Administrative functions
 - All registered templates in Data Xchange and Data Xplorer
 - Access to menu items
 - Job submission privileges
 - Access to the underlying data sets.
-

For more information on the Datalynx DMS and its business application, contact:

Datalynx
Level 5, 92 Pitt St,
Sydney, NSW, 2000

Ph: +61 2 9002 5544

E: sales@datalynx.com.au