A Survey Paper: Content Integration Services

Madhura K

Assistant Professor, CSE Department, Presidency University, Bangalore, India

Abstract: The world of software industry is rapidly evolving. The term enterprise content management was long associated with Organizational practice of using standard set of tools and products housed in data centers for administration of unstructured content throughout its lifecycle, from creation to permanent storage or deletion. However this no longer reflects market dynamics or the organization needs for content in digital business. For IT heads in charge of content management projects, this means casting aside previous notions and rethinking their technology approaches. They should clearly implement in next few years, those technologies that provide consolidated views and analytics across multiple sources, while continuing to deliver engaging and supportive user experiences. This new technology is being called as "Content Integration Services". Integration Services is a platform for building high performance data integration and workflow solutions, including extraction, transformation, and loading (ETL) operations for data warehousing. Integration Services includes graphical tools and wizards for building and debugging packages; tasks for performing workflow functions such as FTP operations, SQL statement execution, and e-mail messaging; data sources and destinations for extracting and loading data; transformations for cleaning, aggregating, merging, and copying data; a management service, the Integration Services service, for administering Integration Services packages; and application programming interfaces (APIs) for programming the Integration Services object model.

Keywords: CM, CIS, ECM, CMIS, ETL, API, CRM

I. Introduction

Content Integration Services are the next stage of enterprise content management, representing a shift from self-contained systems and repositories to open services. Content Integration Services bring in an organization an ability to deliver content services that enable digital business. Enterprise content integration (ECI) is a marketing buzzword for middleware software technology, often used within large organizations, that connects together various types of computer systems that manage documents and digital content. ECI systems often work in tandem with other technologies such as enterprise content management , document management , groupware, and records management. It takes a decentralized approach in order to manage content from various resources. ECI implementations exist on a tier above the organization's existing software and provide wide-ranging search, indexing, and access functions.

A centralized ECM Platform often requires complex integration to deliver information in the context of user activities. A purely centralized platform can also be costly and time-consuming to provide new innovation in fast moving world. Thus realizing the ambition of becoming fully digital. Depending on how an organization has evolved content ends up scattered everywhere, on drives, emails, in business systems, in ECM systems, in on-premise BPM, ERP, BI and CRM system as well. And with advent of SaaS solution content can also be found in the cloud. Technical complexity and in some case vendor lock in due to licensing, makes it very difficult to access all these systems and access right content.

There is a need for Content Integration Services that provides a new way of approaching towards solving future content related problems. Because many organizations now exist in a multi-repository environment, quick wins and solutions that solve problems close to immediately are in demand. That's where Content Integration Services come in that can allow organizations to have more control of various content sources.

Content Integration is multi-disciplinary. It encompasses:

- **Philosophy:** How do we adopt the mindset that content is divorced from its channel? That message and medium are not the same thing, and a message can be carried over multiple media? How do we evangelize this philosophy to the entire organization?
- **Theories:** What are the core paradigms of working with content? What is content, itself? What is a repository? What is a channel?
- **Practices:** How do we design content for integration? How do we manage it in such a way that it can be re-used? What governance and workflow situations arise from the usage of content in multiple locations?
- **Tools:** What type of repository is sufficiently abstracted to allow us to integrate our content easily? What channel products and services are designed for content integration? What content management systems allow for the easy import/export of content for re-use? What systems are easily adapted for use as a Web Content Delivery System?

Content Integration is an umbrella which falls over a collection of knowledge and technology, the combination of which allows us to get more value out of our content – to reach greater numbers of content consumers, at less cost, with greater control, and less risk.

II. LITERATURE SURVEY

Multi-Vendor and selective sourcing have become mainstream in the IT market over the last 10 years. This has led to the need for Service Integration and Management (SIaM) as a distinct discipline that aggregates single IT services, aligning them to core business processes and priorities, as well as presenting them to the end-user community in a consistent "one face to the customer" manner.

Service Integration and Management does not stop at managing internal IT efficiently, When paired with supplementing capabilities such as Cloud Orchestration and Business Process Management, it becomes a foundation for a new level of process agility.

A. *Motivation*

Multi-Vendor Outsourcing: For around the last 10 years, the concept of multi-vendor sourcing has been heavily pushed by advisors and industry experts. There is strong evidence from the market that the majority of medium to large companies across all industry sectors are now following multivendor IT sourcing strategies. The classical way of organizing a multi-vendor environment splits the IT delivery functions into technology towers such as 'data center operations' or 'workplace,' contracting them individually to different outsourcing vendors. The technology tower approach leaves customers' IT organizations with the task of creating and operating an integration and management layer to achieve three main objectives:

- Create end-to-end services that link IT vendor portfolios to ensure overall business requirements are met.
- Define and manage service quality by creating and measuring end-to-end service levels.
- Execute end-to-end support processes by configuration, change, incident and problem management together with other ITIL processes across multiple provider organizations.

B. *Increasing Complexity of Service Management*

It is evident that the higher the number of different vendors that a company has to manage, the more complex and costly this task becomes.

C. *Outsourcing Service Integration*

A rise in service management complexity and costs without a resulting value add in businesses' primary markets, is causing a re-think of the approach to service integration.

III. ARCHITECTURE

 The below diagram shows a reference architecture of a Content Integration Services.
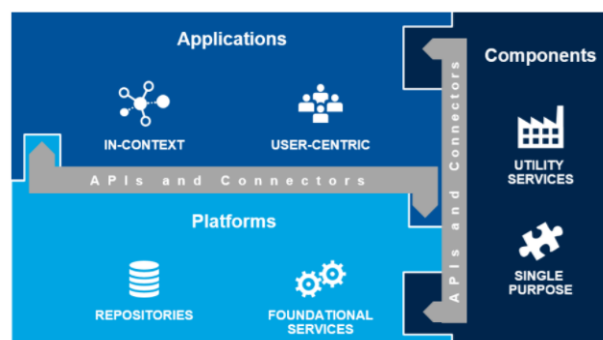

Figure 1.   Content Integration Services Reference Architecture

Content Integration Services consists of –

- **Content Repositories**- Ideally a content repository should be CMIS compliant. CMIS is an open standard for content management interoperability services. This standard provides reference for how a restful service should be provided by CM systems. This CMIS service can then be used by any business applications.

- **Foundational Services**- A CIS should provide a set of foundational services. The services should typically consists the standard functionalities provided by content repositories. In addition this should provide services related to - Regulatory compliance and risk management, Retention and dissemination of business knowledge.
- **API**- A CIS should provide a set of API's in multiple languages and formats like Java, .Net, JQuery, DOJO, REST, WebService.
- **Connectors**- A CIS should provide out of the box connectors to various popular content repositories in market, like connectors for FileNet, Documentum, Sharepoint, Box, DropBox, etc.
- **Authentication and Authorization**- A CIS should provide latest authentication and authorization mechanisms. It should make use of OAuth 2.0, JWT, etc and provide integration to various SSO products in market with support on SAML, O365 integration.

IV. ADVANTAGES

A Content Integration Services provides multiple advantages to an organization portfolio-

- It can support regulatory compliance by supporting traditional on premise content repositories meant to handle such requirements. Furthermore with stringent rules on data classification and storage requirements, it makes perfect sense to have a product that can connect to multiple repositories to meet the requirement.
- It provides Risk management capabilities by providing safe access to original repositories.
- Retention and dissemination of business knowledge from various channels and sources, like mainframe, content management products, ERP products, etc. Information is extracted from multiple source systems and distributed to multiple destination systems.
- Cost and process efficiencies.
- Innovation and new ways of working.
- Rapid time to market.

V. EXISTING STATUS

A. Tools and Utilities

- This section describes how to use Business Intelligence Development Studio and SQL Server Management Studio with Integration Services and the tools and utilities that Integration Services provides.
- **Integration Services and SQL Server Management Studio**.- Describes how to use Business Intelligence Development Studio and SQL Server Management Studio with Integration Services.
- **Integration Services in Business Intelligence development Studio** -Describes all the features that are available in Business Intelligence Development Studio for developers of Integration Services packages.
- **SSIS Designer**- Describes the graphical tool for building complex packages that include control flows, data flows, and event-driven logic.
- **Integration Services Wizards-**Describes the wizards for importing and exporting data, installing packages, migrating packages, and creating package configurations.
- **Command Prompt Utilities-** Describes the Integration Services command prompt utilities for managing and running packages.
- **Query Builder Overview -**Describes the graphical tool for building queries.
- **Expression Builder-**Describes the graphical tool for building expressions.

B. Content Management Interoperability Standard

Modern Content Management tools are using CMIS standard to evolve their products into content integrations services.

Content Management Interoperability Services (CMIS) is an open standard that allows different content management systems to inter-operate over the Internet. Specifically, CMIS defines an abstraction layer for controlling diverse document management systems and repositories using web protocols.

CMIS is an OASIS standard. CMIS standard defines a domain model and Web Services and Restful AtomPub bindings that can be used by applications to work with one or more Content Management repositories/systems.

The CMIS interface is designed to be layered on top of existing Content Management systems and their existing programmatic interfaces. It is not intended to prescribe how specific features should be implemented within those CM systems, not to exhaustively expose all of the CM system's capabilities through the CMIS interfaces. Rather, it is

intended to define a generic/universal set of capabilities provided by a CM system and a set of services for working with those capabilities.

C. *Products available in the market*

- **SharePoint-** Microsoft's SharePoint offerings including SharePoint Online, SharePoint Server, Office 365, OneDrive for Business as well as integrations with Teams and Flow in Office 365 and other extensions to make it a true CIS.
- **OpenText-** OpenText provides two Content Integration Services, its original OpenText content server and Documentum provides CMIS capabilities.
- **IBM**- is providing CMIS based FileNet platform. Along with this it's integration with BOX and IBM Content Navigator capability to integrate multiple content repositories makes it a formidable CIS provider.
- **Xillio**- Xillio is a true CIS product, allowing integrations to almost all content repositories. It also provides a transformation feature that can transform metadata and detect and remove de-duplication in content and metadata.
- **Vega**- Vega is also a CIS product allowing integrations to almost all content repositories. However it doesn't provide any UI support and rely on native products UI or business application to provide UI to users.

VI. CONCLUSION

Modern organizations are rapidly moving towards adaption of new technologies. With the help of content integration services, organizations can become truly digital. However without any checks and balances IT spend can drastically go out of control, having to manage multiple licenses and service management costs. This is truly a double edged sword. Each organizations should develop a standard unified content policy that govern and directs the strategy towards content integration. Furthermore organizations should study and invest in a robust content integration service product. This will allow them to become fast pacing and keep up with rest of the world of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. PLEASE DO NOT RE-ADJUST THESE MARGINS. Some components, such as multi-leveled equations and graphics, are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

REFERENCES

[1] M. Sumner, "Risk factors in enterprise wide information management systems projects", SIGCPR '00: proceedings of the 2000 ACM SIGCPR conference on Computer personnel research, ACM, pp. 180 187, 2000.
[2] E.Bertin, P. Lesieur, "Which architecture for integrated services?", International Networking and Services, 2006.
[3] Functional Description Release 1, June 2009.
[4] T. Abe, "What is Service Science" in, Tokyo: Fujitsu Research Institute, Economic Research Center, Dec. 2005.
[5] Martin Gebauer, Fred Stefan, "Integration Services Today: A Qualitative Study -- Design and Preliminary Results", Web Services (ECOWS), 2011 Ninth IEEE European Conference on WebServices.
[6] http://docs.oasis-open.org/cmis/CMIS/v1.0/cmis-spec-v1.0.html
[7] https://blogs.gartner.com/michael-woodbridge/the-death-of-ecm-and-birth-of-content-services/
[8] https://www.xillio.com/
[9] https://products.office.com/en-us/sharepoint/sharepoint-online-collaboration-software
[10] https://www.ibm.com/us-en/marketplace/content-navigator
[11] https://www.ibm.com/us-en/marketplace/case-foundation
[12] https://www.vegaecm.com/migrations/
[13] https://www.opentext.com/what-we-do/products/enterprise-content-management/extended-ecm-
[14] https://gadgetopia.com/post/9272
[15] https://technet.microsoft.com/en-us/library/ms137812(v=sql.105).aspx
[16] https://atos.net/wp-content/uploads/2016/07Atos-Ascent-Wite-paper-Service-Integration-and-Manaement-Motivation-Challenges-and-Best-Practices.pdf