Integration Platforms – Problems and Possibilities[*]

Hristina Daskalova, Tatiana Atanassova

Institute of Information Technologies, 1113 Sofia

Abstract: *The purpose of the paper is to investigate the specifics in the development of the integration technologies, connected with the uniqueness of the existing information systems and the necessity for their interaction when new ideas, objectives and tasks are presented by the information resources users. They are pretty new technologies, so not all of the basic standards have been definitely discussed, formulated and accepted, which might cause certain risks in the development of integration projects. The present paper proposes (without any discussion of the technical details) a short review and analysis of the problems, connected with the information resources integration, the integration platforms available, the characteristics of the data integration, the processes, the applications and the interfaces, as well as the applicability of integration in service-oriented architecture. An exemplary technical model of an integration platform is suggested. A large number of the topics discussed are rather complex from a technical viewpoint, so many papers, books and technical documents could be found for each one of them. The authors have focused their attention on the question why the research work in the field of information resources integration and the design of integration platforms is considered as one of the challenges in information technologies development.*

Keywords: *integration, information resources, service-oriented architecture.*

1. Introduction

The great variety of software products, existing nowadays, has overlapping functionality and incomplete compatibility, because their functional possibilities rather, not their interaction with other components has been accentuated in their

---

development. The increasing necessity for solution of problems, connected with the information resources integration and the control of the information exchange provokes consideration of the following:

– the heterogeneousness of the information stores,

– the increasing activity of the information stores contents,

– the complete character of the services with respect to contents storing and supply,

– the enhancing necessity for new functionality and facilitation of the information access.

In case no integration tools are used, manual encoding is necessary, which is pliable to errors, without any optimal completion, and leads to waste of time, resources and impossibility for repeatable use.

The investigations are directed towards *automation* of the integration with the purpose to reduce the time and the resources, necessary to achieve efficient solutions in the work of enterprises and organizations. In the future the user will just declare what information he needs, its layout and properties, and the information system (in the ideal case) will answer the request, employing the user in the control of some key solutions only [1]. The information will enter from heterogeneous resources and it is not expected that the user will participate in the selection of the integration technology.

2. Integration of the information resources

The integration of the information resources aims at ensuring the integrity and non-discrepancy of the available information, avoiding the input data duplication, improving the information exchange and the efficiency of the entire information infrastructure, as well as the creation of a uniform information space.

The integration could be accomplished at several levels:

– at data level;

– at services level;

– at interfaces level;

– at applications or business processes level.

Besides certain functionality, the integration of the information resources requires the construction of an infrastructure, providing safety and security.

This imposes the development and use of *specialized integration platforms*, that are a software complex with wide functional options, consolidating the logic of the solutions to given integration problems within the frames of the whole information system.

3. Integration platforms

The integration platforms enable the applications interaction, attempting to leave them unaltered, the integration functions being separated at that, thus achieving a possibility for flexible alteration of the integration logic in the platform itself, without any considerable influence on the system integrated.

Definite developments and choreography of the applications work lead to: formalization and automation of the sequence, defined by the real business processes; providing of the connection between the different components – applications; transformation of messages formats; differentiation between the logic of the business process and the logic of the applications.

Depending on the character of the integration problems, the integration platforms could be divided as follows:

– platforms for *data* integration (consolidation and synchronization of the data bases);

– platforms, providing the interaction between different *applications* (data exchange, invoking of certain functions);

– platforms for integration of business *processes* (a possibility is provided for complex automation of the business processes);

– platforms-portals, which accomplish consolidation of the user's access with the help of a common point of access to the information resources.

An exemplary technical model of an integration platform is given in Fig. 1.
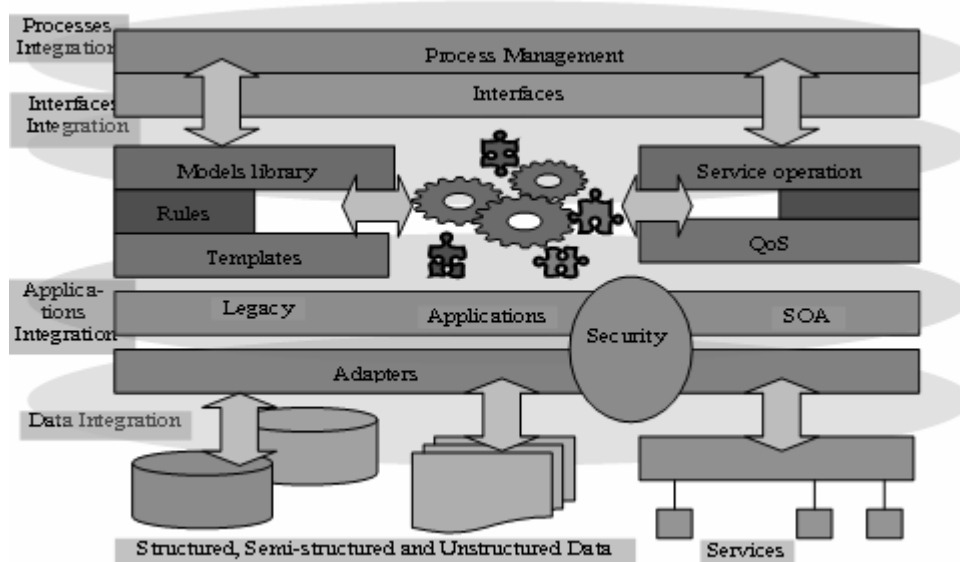

Fig. 1. An exemplary technical model of an integration platform

4. Data integration

The approaches for consolidation of the data from the information resources include:

- modeling of the resources, which plays a central role in applications integration and interoperability of the business processes;
- data stores (Data Mart/Data Warehouse/Meta Data Repository);

- standards for internal and external applications – Web Ontology Language (OWL), Resource Description Framework (RDF) and Extensible Metadata Platform (XMP);
  - processing of both structured and unstructured data;
  - methods for search in heterogeneous resources.

The following standardized integration technologies may be pointed out:

EAI – Enterprise application integration – a system of methods and tools for consolidation and co-ordination of the applications in enterprises [5];

EII – Enterprise information integration, which enables the integration of data from different resources. This technology is useful in the design of one point and one language of access, providing virtual integration of the data from different resources [4];

ETL – Extract, transform and load – these are processes, which accomplish the transfer, re-formatting and deletion of the data from combined resources (data bases or data warehouse), the synchronization of the data bases, analysis and insurance of the business processes [6].

These integration technologies can be used to create a homogeneous infrastructure of data that enables the discovery, transformation and use of the information required.

5. Applications integration

The main differences among the applications are determined by their architectural components, such as the data model, the technologic realization (basic software, environment and time of execution, server used), and also by the model of the business process. The differences in business processes models, as well as the mechanisms for their realization, do not allow the acceptance of the existing applications as part of a global business process. Besides this, the applications themselves are realized as autonomous systems, which are not a part of any global process. Regardless of the internal architecture of the applications, they must have Application Programming Interface (API), published and legally used, in order to accomplish an accessible operation. An indicator can be searched for, characterizing to some extent the capacity of a given application to be a part of a global (or composite) application. The applications are altered, modified, re-processed, new versions are created and parallel with this the programming interface is changed and this must be reflected in its publishing. The evaluation of the applications integration degree is still more complicated by the applications' authors, who do not wish to reveal the internal design of their applications, attempting to keep the copyright on the software product developed. The situation with the industrial applications is quite different, since their authors usually design some adequate programming interfaces, but often use a programming language, which is accessible to a given user only. The development of integration projects supposes that different companies develop adaptors for programming access to the industrial applications without any specific information about the programming interfaces.

6. Interfaces integration

Another way to solve the integration problems is to design interfaces that provide interaction of "point-to-point" type of the applications, the integration applications being united at that. Such is, for example, the approach of remote procedures (Remote Procedure Call, RPC) in the applications "client-server". The main advantage of interfaces integration is the preservation of the applications already existing, realizing the possibility for information exchange with the help of interfaces with apriori given interaction rules. The following difficulties could be mentioned for this integration approach: at a larger number of the applications, participating in the interaction process, the number of the connections increases as well, which complicates the architecture of the information system and hampers its exploitation and development; the organizing of the interaction with external information systems requires data transfer over public networks and their protection, which complicates and restricts the possibility for integration; during the integration process each application must contain information of the applications addresses it communicates with, and it must also be able to transform data formats.

In this way the concordance of the data formats, the messages routing, the ensuring of reliability and security are solved separately for every pair of messages. Even for simpler interaction scenarios the integration logic becomes more complicated. In addition, the interfaces integration is accompanied by the development and change of the existing programming code in one or in both applications being integrated.

The realization of the interaction of a new application with an already existing one requires the formation of new connections. Taking into account the characteristics of this approach, it is obvious that its usage is appropriate in the integration of no more than three applications, which have similar formats of data storing and use identical technologic solutions. The architecture of the information system, its maintenance and possibility for development are rather complicated in this approach.

7. Processes integration

The integration of the processes is a high level of the business integration. It includes the co-ordination and control of the operations, which may involve numerous systems and users' roles. This type of integration ensures the structuring, realization, automation and control of the business processes with simultaneous measuring of the parameters necessary for subsequent optimization of the processes models. The processes integration supports long term transactions and operations with users' participation (on the basis of their roles). The flow of the business events, entering a process can be modified with the help of external signals – on the basis of some parameters, determined during processes adjustment, or accounting the data from external resources, like the application data base, or on the basis of users' solutions.

The processes integration may be regarded as a level of business logic, which determines what must be done at each step of the process, in contrast to the way of process realization, which is defined within the application itself. The separation of the logic, determining the choice of operations, from the tools for their accomplishment, ensures additional advantages of the processes, since the changes in the realization of the two stages can be done independently.

8. Service-oriented architecture

The service-oriented architecture allows step by step integration. It is based on the earlier approaches towards integration – data sharing and Remote Procedure Call (RPC), which enable the introduction of the concepts "service provider" and "service consumer", "platform independence" and "connectivity".

The idea to use Service Oriented Architecture (SOA) suggests the development of composite applications upon the services as components [7, 12]. The services in the composite application inter-integrate and the integration platform becomes a part of each information system

The accent is on the functional aspect of integration and the adoption of service-oriented architecture that allows the creation of modified business services and the reorganization of the business processes with the help of some developed modules.

SOA technology, which has appeared as a successful approach to programming and remote procedure call in Internet, is nowadays applied in the applications integration. The intense investigations attempt to expand the service-oriented technologies by semantic models and methods.

Some companies, such as SAP, IBM, Microsoft and Oracle determine the new tendencies on information technologies market, investing in the development of integration platforms. The latter include not only technologies at a transport level, ensuring the necessary asynchronous connection among the applications, but also the possibility for complex transformations of the information, transmitted between the applications.

The progress of the integration technologies shows a tendency towards development of the functionality and standardization. The platform for programming J2EE supports mainly the standards, referring to SOA (BPEL Business Process Execution Language for Web Services, SDO - Service Data Objects, SCA - Service Component Architecture). Together with this, some inter-platform standards are developed in Web-services area.

The integration platforms SAP NetWeaver, Oracle Fusion MiddleWare, IBM WebSphere and a series of Microsoft servers are most widely spread in practice.

9. SAP NetWeaver – the complete integration
and application platform

The advance of open technologies has intensified the competition between the business leaders. SAP has answered the challenge by SAP NetWeaver – a complete

open integration platform for applications. SAP NetWeaver operates with the existing IT infrastructure, accounting and controlling the altering conditions [10].

SAP NetWeaver platform (particularly SAP NetWeaver 2004) has an expanded set, including a united portal, support of mobile devices, Web-interface for users, etc. The suggested SAP AG architecture – Enterprise Services Architecture (ESA) has a distributed software model. Independent Web services are used for the program realization of the business processes that can be with multiple usages in combinations, determined by business logic. The main components of the integration platform are: SAP Enterprise Portal, SAP Business Intelligence, SAP Exchange Infrastructure, SAP Application Server, SAP Master Data Management and SAP Mobile Infrastructure. The implementation is done in parts – by different platform components, the entire use of SAP NetWeaver is also expected. SAP NetWeaver Portal, considered as a central element of the information system, that enables the joint work of users through Web-interface with controlled access to co-operative and general information, is quite popular. The technology of users' single authentication is realized in the portal and common co-operative safety policy is applied, which simplifies the interaction of already operating software products and data bases. SAP NetWeaver Exchange Infrastructure component removes the duplication of the procedures in data entry, realizes control on the operations and accelerates requests processing.

With SAP NetWeaver the design, construction and application of new business strategies is simplified. SAP NetWeaver aids the control of innovations in the organizations, combining already existing systems. SAP NetWeaver integrates people, information and business processes, which use different technologies in and outside the organization, controlling structured and non-structured information, harmonizing the data, earlier supported in different systems with the help of the solutions for business analyses, knowledge control and general data control. Besides this, SAP NetWeaver supports some Internet standards, such as HTTP, XML and Web services, as well as openness and interaction with Microsoft.NET and J2EE – particularly with IBM WebSphere. In this way it takes advantage from the best solutions of business control. As an integral applied platform SAP NetWeaver preserves and provides the benefits from the already made IT investments and gives the bases for innovatory and flexible inter-company processes, typical for the future. SAP NetWeaver is the basis of SAP xApps and of the solutions of mySAP Business Suite family. It lies in the basis of SAP Enterprise Services Architecture, which combines the experience of SAP in the area of business applications with the flexibility of the Web services and open technologies – for integral business solutions in the form of services.

10. Integration platform of Oracle Fusion Middlewares integration solution

The organizations realize the benefits of SOA architectures implementation and often seek for software like Oracle SOA Suite and Oracle Business Process Analysis Suite, key components of Oracle Fusion Middleware. Designed on open

standards and architecture, enabling the upgrading of the system without interrupting the systems already existing (the so called hot-pluggable architectures), Oracle SOA Suite package [11] facilitates the integration of applications and data in heterogeneous environments. The clients have the opportunity to model, control, protect and develop web-based services for business processes, maximizing in this way the benefits from their business and protecting their investments. Fusion Middlewares Integration solution of Oracle is a powerful and reliable integration platform, which enables the reduction of the business expenses for software purchase and the optimization of already available applications by consolidation and optimization of the business processes. Oracle Fusion Middleware offers the clients an integral platform-independent SOA and BPM software, which helps the raising of co-operative flexibility simultaneously with the decrease of the expenses. Suggesting one reliable SOA infrastructure, Oracle Fusion Middleware enables the clients to protect the investments already put and also their increase in the direction of the following generation of architectures. Forrester gives the results from a survey "The Forrester Wave: Integration-Centric Business Process Management Suites" [2], devoted to packages for integrated control of the business processes. The analysts define Oracle as a leader and give high evaluation of the key components of Oracle® Fusion Middleware – a family of standardized solutions for development that allows the design and control of SOA services by the users in heterogeneous computing environments. The survey is based on 85 criteria, covering the strategy and market activity and makes the conclusion that the possibilities of Oracle SOA Suite – a strategic component of Fusion Middleware, guarantee the support of Service-Oriented Architecture − SOA and the control of production processes (Business Process Management − BPM). The analysts consider in details the new functional possibilities of Oracle SOA Suite, both in processes modeling and also in the control of the business activities, the working flow and in suggesting an improved support of the business rules. Oracle expands also the possibilities for access to its platform thanks to the fact that Oracle SOA Suite can work in other J2EE platforms too.

11. Integration solutions of IBM WebSphere

The main strategy of IBM is to maintain an open heterogeneous platform, including different data bases, portals, Internet-servers and content stores. IBM WebSphere program solution enables the flexible integration of different applications on various platforms, uniting infrastructural and integration software tools, designed to create environment that is flexible with respect to the alterations; provides functional reliability and increases the productivity of the business processes [8].

The fundament, used to build IBM WebSphere, is SOA. The service-oriented architecture enables the presenting of the problem in the form of separate realizable functional parts, with no necessity for surplus investments in information infrastructure development. In order to ensure complex integration at all the stages – modeling, design, development and control, IBM provides a number of products, supporting SOA, which are divided in four basic groups: IBM WebSphere Business

Modeler, WebSphere Integration Developer, WebSphere Enterprise Service Bus (ESB), WebSphere Business Monitor (Business Process Management) (Fig. 2).
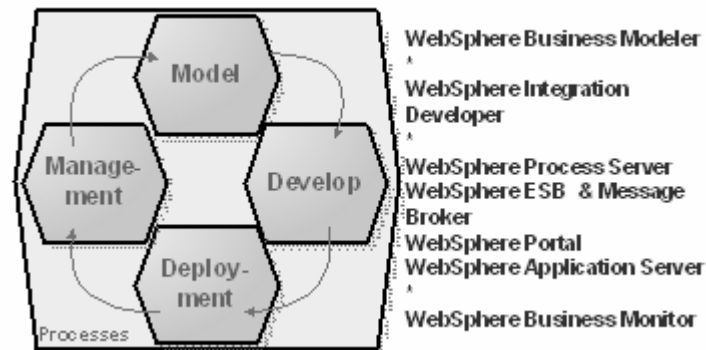

Fig. 2. IBM WebSphere platform

WebSphere Business Modeler – a tool on the basis of Eclipse, which enables the graphic modeling of the business processes in BPEL or FDL format.

WebSphere Integration Developer allows the development of some composite applications, regarding the existing IT resources as services, which could be interconnected in complete business processes.

Enterprise Service Bus (ESB) that is a flexible communication infrastructure for integration of the applications and services, expands SOA capacities, reducing the number and complexity of the interfaces among the applications and services. ESB does the following:

- it distributes the messages among the services;
- converts the transport protocols between a notifier and a service
- converts the format of messages between a notifier and a service;
- controls business events from different sources.

For ESB IBM offers two basic products: IBM WebSphere ESB and IBM WebSphere Message Broker. WebSphere Message Broker is designed to realize unified communication interaction and transformation of different messages formats (XML data).

IBM WebSphere Business Monitor enables the control of the business processes in real time, visualizing their status.

A number of products, satisfying different business requirements, are developed in addition to these tools. IBM Mashup Starter Kit enables the construction of mashup on the basis of Web 2.0. The kit comprises two key technologies. IBM Mashup Hub is a mashup-server, that supports the information flows in RSS, ATOM or XML formats for their multiple and combined use. Mashup Hub enables the uniting, transform, filtering, annotation and publishing of the information in new formats. The other component – QEDWiki serves as user's interface to derive information from an arbitrary data source and to create promptly an uniform representation of different types of information.

IBM WebSphere Process Server includes a set of functions for maintenance of short term or long term processes.

SOA Security and Management configuration suggests some patterns for solutions and development of a reliable infrastructure for protected and controllable SOA environment. IBM Optim new solutions realize some means for data control as a component in SOA environment.

The expanded professional services in the control areas of SOA environment include the design and planning of service management, the introduction of tools controlling the infrastructure for administering and monitoring, as well as services of the test center.

IBM WebSphere Application Server Feature Pack for Web 2.0 enables the connection with external Web-services, internal SOA-services and J2EE objects with the purpose to create interactive interfaces to Web-applications.

12. Discussion

The considered integration platforms aim at the realization of flexible environment for the design, development and execution of transactions and business applications.

Unlike some earlier developments, which have ensured local-based interaction among the software modules, realized in different languages, the modern solutions are oriented towards network infrastructure. In global network interaction the different modules are independent and distant, they do not use a uniform address field and the processes do not run on one and the same computer. The communication channels for transmitting messages and protocols during the communication among the software components are also objects of integration in the network infrastructure while combining the information resources. The examined solutions applied in the integration platforms are based on distributed service-oriented architecture – SOA for creating a unified way for communication.

The system integration methods on the basis of SOA and XML-specifications provide a flexible mechanism of access, facilitating the process of combining different network applications in the interaction, functioning as a whole. In this way all the above mentioned software manufacturers enable flexibility in the integration of the information resources and overcome the detachment of the business applications.

In addition, the service-oriented architecture, used by SAP, IBM and Oracle, supports an incremental approach in the design of integrated information systems. In such a way SOA and the modular principle enable the successive integration of the components and facilitate the verification and validation of the integration solution.

The standardization of the infrastructure – networks, servers, data bases, configuration components, could facilitate the uniting of heterogeneous components. The companies-leaders above mentioned are working intensely in this direction.

SAP, Oracle and IBM realize the interoperability by utilizing their existing software products, upgrading them and adding new modules and components. The interoperability of the applications is expressed with the help of functional, syntax and semantic methods. In order the information to be machine processed by the receiver and shared among the system components, it must be semantically interoperable, i.e. understandable at the level of the formally defined concepts in the applied area. Up to now, this is still in the area of investigations and experimental studies, but also of some industrially introduced solutions.

The use of SOA together with weakly connected XML-based web services, standards, *open* technologies and API for creating of modularity that provides of geographic, server, network distribution and reliability for the requirements of business – these are the characteristics of modern integration platforms.

13. Conclusion

The integration of the information resources remains a hard problem. It includes the connection of different user's applications (C++, Java, J2EE), complex solutions (such as SAP CRM, ERP) and functional programms (such as IBM Information Management System [9]), that may be geographically distributed and may be implemented on heterogeneous platforms.

The lack of standard methods for integration of the information resources [3], technologies and business processes hampers the supply and the efficient use of the information by the users. It is necessary to clarify and develop the norm recommendations, in order to define the requirements towards the numerous *components* that will be connected to the integration platforms:

- Web-services;
- integrated data;
- metadata of the content;
- identifiers;
- computers for access;
- limited access channels;
- mobile telephones for access;
- systems for video conferences with the help of an IP-protocol;
- systems for voice exchange systems with the help of an IP-protocol;
- smart cards for access;
- Business processes and others...

The methods for integration themselves, the tools for standardization, specification, semantic description of the resources, as well as the corresponding theories is still to be developed.

In order to achieve this objective, it is necessary:

- to raise the level of abstraction in information presentation, i.e. presenting the logic of its inter-connectivity in machine comprehensible form;
- to enrich the information by meta data, automatic extraction of the meta data;

- to automatically supply the entire necessary information, i.e. automatic recognition, search, discovery of the needed data from different resources.

No fundamental theory for information resources integration has been offered until now. The semantic description and realization of some methods for semantic consideration may be the key to the construction not only of some isolated solutions, but also to the achievement of common integration objectives.

The integration of the information resources prepares the infrastructure for realization of service-oriented architecture (**SOA),** which stimulates the general development of the information structure, involving in the integration layer some new applications, not setting at that any constraints on the used technologies and platforms.

R e f e r e n c e s

1. H a a s, L a u r a. Beauty and the Beast: The Theory and Practice of Information Integration. – In: 11th International Conference on Database Theory (ICDT 2007), Barcelona, Spain. January 10-12, 2007.
   http://www.almaden.ibm.com/cs/people/laura/ICDT_Haas.pdf
2. The Forrester Wave: Integration-Centric Business Process Management Suites. 2006.
   http://www.oracle.com/corporate/analyst/reports/infrastructure/fm/Forrester-BPMS-Q42006.pdf
3. R u s s e l l R u g g i e r. Integration Theory. Part 1. DM Direct, August 2005.
4. I m h o f f, C l a u d i a. Understanding the Three E's of Integration. – DM Review, April 2005.
   http://www.dmreview.com/article_sub.cfm?articleId=1023893
5. EAI definition.
   http://searchwebservices.techtarget.com/sDefinition/0,,sid26_gci213523,00.html
6. ETL definition.
   http://www.computerworld.com/databasetopics/businessintelligence/datawarehouse/story/0,10801,89534,00.html
7. B r o o k s, R e x and R u g g i e r o, R u s s e l l. EA & SOA. – DM Direct, April 2005.
   http://www.dmreview.com/article_sub.cfm?articleId=1025078
8. IBM Websphere Software.
   http://ibm.com/software/websphere
9. IBM Information Management.
   ibm.com/software/data/
10. SAP NetWeaver.
    www.sap.com/solutions/netweaver
11. Oracle SOA Suite.
    http://www.oracle.com/technology/tech/soa/index.html
12. SOA.
    http://webservices.xml.com/pub/a/ws/2003/09/30/soa.html