http://doi.org/10.35784/iapgos.3295 *received: 15.11.2022 | revised: 13.01.2023 | accepted: 12.03.2023 | available online: 31.03.2023*

DIGITALIZATION OF ENTERPRISE WITH ENSURING STABILITY AND RELIABILITY

Gulnar Balakayeva[1], Paul Ezhichelvan[2], Yerlan Makashev[1], Chris Phillips[2], Dauren Darkenbayev[1], Kalamkas Nurlybayeva[1]
[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan, [2]Newcastle University, Newcastle, United Kingdom

Abstract. *The article is devoted to the development of an information system for automating business processes of a modern enterprise with ensuring stability and reliability, which are implemented by the applications developed by the authors. Goal is to develop improvements to the core digitalization processes of enterprises for sustainable functioning. The authors carried out a deep analysis and described the main stages of the enterprise digitalization process: the process of document approval, business processes of personnel management, etc. The architecture of the information system, a description of business processes and the principles of reliability and fault tolerance of the system being developed have been developed. The developed desktop-client application provides connection to the information system with the help of working computers of the enterprise through a local network with access to the application server. This allows you to reduce damage from accidental or deliberate incorrect actions of users and administrators; separation of protection; a variety of means of protection; simplicity and manageability of the information system and its security system.*

Keywords: digitalization, business processes, system architecture, reliability, data security

CYFRYZACJA PRZEDSIĘBIORSTWA Z ZAPEWNIENIEM STABILNOŚCI I NIEZAWODNOŚCI

Streszczenie. *Artykuł poświęcony jest opracowaniu systemu informatycznego do automatyzacji procesów biznesowych nowoczesnego przedsiębiorstwa z zapewnieniem stabilności i niezawodności, które są realizowane przez opracowane przez autorów aplikacje. Celem jest rozwijanie usprawnień podstawowych procesów cyfryzacji przedsiębiorstw dla zrównoważonego funkcjonowania. Autorzy przeprowadzili dogłębną analizę i opisali główne etapy procesu cyfryzacji przedsiębiorstwa: proces akceptacji dokumentów, procesy biznesowe zarządzania personelem itp. Zostały opracowane architektura systemu informatycznego, opis procesów biznesowych oraz zasady niezawodności i odporności na błędy tworzonego systemu. Opracowana aplikacja typu desktop-client zapewnia połączenie z systemem informacyjnym za pomocą pracujących komputerów przedsiębiorstwa poprzez sieć lokalną z dostępem do serwera aplikacji. Pozwala to na ograniczenie szkód wynikających z przypadkowych lub celowych nieprawidłowych działań użytkowników i administratorów; rozdzielenie ochrony; różnorodność środków ochrony; prostotę i łatwość zarządzania systemem informatycznym i jego systemu zabezpieczeń.*

Słowa kluczowe: digitalizacja, procesy biznesowe, architektura systemu, niezawodność, bezpieczeństwo danych

Introduction

The system architecture supports a transactional model that guarantees the integrity of system data throughout all stages of their life cycle. Automation system of the enterprise information system provides automation of formation and accounting (registration) of administrative documents, memos, meetings, and incoming and outgoing correspondence. Clients working outside the enterprise can access the information system through a web browser, as a web client, via a web server. The client application has two possible types of connection to the enterprise information system: a desktop client and a web client [2].

The proposed information system is designed for an efficient operation of the administration (user, client) of the enterprise [3]. The client application has two possible types of connection to the enterprise information system: a desktop client and a web client [19]. The desktop-client application enables connection to the information system using the working computers of the enterprise through the local network, with an appeal to an application server. Clients working outside the enterprise can access the information system through a web browser, as a web client, via a web server [16].

The proposed information system has a layered architecture using a domain-specific tool. The architecture incorporates provisions that seek to guarantee scalability, reliability and security of the system as we explain later [4]. By scalability we mean that the system should be able to handle a potentially large number of users without serious impact on system throughput and response times; reliability provisions would include mechanisms for tolerating crashes of system components with which the users would directly interact; finally, security considerations will be an integral part of the architecture, cutting across all levels of the system from the frontline clients tier to the database backend tier [13].

1. Multilevel architecture of the enterprise information system

The proposed information system has a layered architecture using a domain-specific tool. The architecture is the guarantor of the availability, reliability and security of the system [14].

Client applications – applications for end-users, development tools, system administration utilities. The client can be either a Windows application (inside the enterprise) or a web browser (client outside the enterprise).

Server components (service management agent) – application servers host a variety of services that allow clients to perform document processing services such as storing new documents in a Database Management System (DBMS) (backend or data tier), fetching relevant documents from the DBMS and modifying them as per application needs and storing the modified versions back in the DBMS. Supporting these services form the core functionality of the information system. These services fall under two broad categories: general administration and human-resource related [11].

This architecture proposes that these core services be instantiated on more than one host for reasons of scalability and reliability. Thus, a client can connect to any one of these hosts to conduct their document processing activities. Typically, an incoming client is exposed to a list of available hosts from which it randomly selects one host to connect to. Thus, if there are N clients simultaneously accessing the system and if there are H (>1), server hosts, then each host is expected to be servicing, on average, only N/H clients at any given time. Thus, by increasing H as N increases (e.g., during peak usage hours) the average load handled by each server is kept constant, and thereby the overall system performance as seen by clients is maintained [12].

DBMS – storage of data and metadata of documents related to the enterprise. It may also be referred to as the backend or data tier.

File storages (data warehouse) – archives of large or rarely used documents, which are more efficient to keep outside the DBMS; these are managed by their own services.

As indicated above, being part of an organisation's information infrastructure, the system architecture needs to demonstrate certain characteristics that are important for any corporate system:

Reliability. The system architecture supports a transactional model that guarantees the integrity of system data throughout all stages of their life cycle. File document storages allow one to organize reliable storage of documents.

Security. For each object within the system, it can be specified which users or groups have the right to perform certain actions with it. Confidential electronic documents and tasks can be encrypted directly in the system by any Microsoft CryptoAPI-compatible encryption provider, which guarantees protection even from users with unrestricted access to data. Logging of all user actions will allow restoring the history of work with system objects in the event of a security violation. This provides high protection against unauthorized access to document storages of all types.

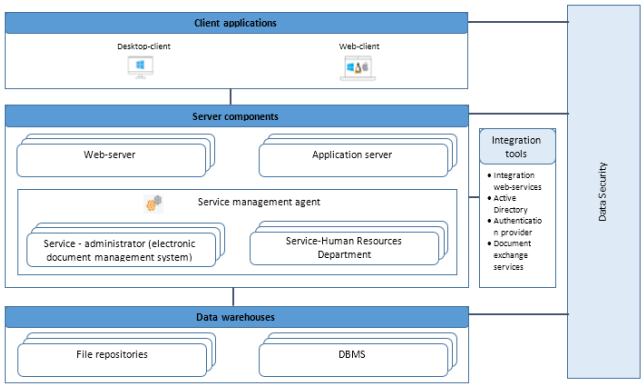

Fig. 1. Enterprise information system architecture

2. Enterprise digitalization system. Document processing

This supports the storage, processing and management of electronic documents and is designed to simplify the task of document processing by clients in a secure, scalable and reliable manner [17].

For a comprehensive solution to the problems of managing electronic document processes, the system must address the following requirements:

- The digitalization system of the enterprise (Fig. 2) makes it possible comprehensively to automate any work associated with contracts and related documents, including electronic approval. This provides the formation of documents by templates, the formation of registration numbers, registration of storage locations, registration of the availability of original documents, etc.
- Automation system of the enterprise information system provides automation of formation and accounting (registration) of administrative documents, memos, meetings, and incoming and outgoing correspondence.
- Electronic document management system supports the automatic approval of documents, the formation of mailings of documents for review, provides routing of documents, etc.
- System of control over execution of documents and orders allows the automatic formation of orders for the execution of documents, and provides automated control over deadlines.
- Electronic archive of documents solves the problem of archival storage of electronic documents. It allows you to create an archive structure from simple and moderated folders, and supports a search for documents in the information storage database
- System of notifications and reminders provides automatic and timely notification of process participants about current and upcoming events, and allows employees to exchange messages.
- Business process automation system is the core of the system. It is intended for the formation of graphic models of automated processes and ensures the movement of documents along routes. It manages a dynamic role-based process access model. Interacts with all subsystems.

When creating complex information systems, designing their architecture, infrastructure, choosing components and connections between them, a number of specific conceptual requirements aimed at ensuring the security of functioning are taken into account, in addition to the general ones (openness, scalability, portability, mobility, investment protection, etc.) [15]:

- The system architecture is flexible enough i.e. it allows relatively simple development of the infrastructure without fundamental structural changes, and changes in the configuration of the used means, and an increase in the functions and resources of the Information System (IS) in accordance with the expansion of the spheres and tasks of its application;
- The security of the system will be ensured under various types of threats and reliable data protection from design errors, destruction or loss of information, as well as user authorization, workload management, backup of data and computing resources, the fast restoration of IS functioning;
- Will provide comfortable, maximally simplified user access to the services and results of the functioning of the IS based on modern graphical tools, mnemonic diagrams and visual user interfaces;
- The system will be accompanied by updated, complete documentation, providing qualified maintenance and the possibility of developing the IS.

It is proposed to do processing documents according to the following schemes.

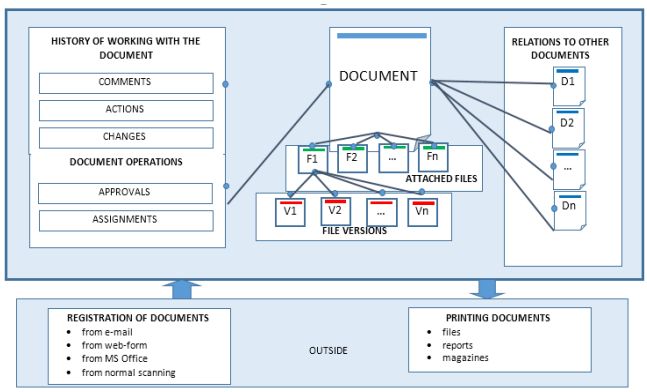

Fig. 2. Enterprise digitalization system

3. Document approval process

Document approval is the most often demanded business process in any organization. The main purpose of the business process for the organization is to control and monitor the terms (dates) of approvals.

Most companies face the task of approving a document [18].

The steps in the document approval process are suggested as follows:

1. Start of the process. Approval of the document with the indication of the document type -> The initiator fills in the main form for approval
2. Preparation of the document -> form is filled with data according to the template for the document type (for references -> number, executor, text of the reference, from (who), to (whom), date of the execution, for contracts -> contract number, contract date, text of the contract, the electronic version of the document, etc.). According to the type of document, the period for which the participants

in the process must agree on the document is determined automatically: agreement – 5 working days, order – 3 working days, order – 1 working day, memo – 6 working days, etc.
3. Identification of direct supervisors of the initiator for approval. The document is assigned "On approval by the head" status. The head of the initiator either rejects the contract by adding comments or agrees. In the first case, the contract is sent to the initiator for revision, in the second case it goes to the next stage and receives the status "Under agreement with the executers".
4. Approval of the document by the head of the executor's department. The performer is selected. In parallel for all performers the execution sub process is started.
5. After the execution of all participants, the result is collected.
6. Notification of managers about the execution.
7. Document registration. Assigning a number according to the classifier and the nomenclature of the organization's affairs.
8. Completion of the process.

At each stage, it is possible to download, print and return the document for revision.

4. Data security architecture development

Only a proven architecture is able to manage effective merging of services, ensure the manageability of the information system, and is able to evolve and resist new threats while maintaining properties such as high performance, simplicity and usability [1].

From a practical point of view, to ensure security, the following principles of building an IS architecture (Fig. 1) are most important:

1. Designing the IS based on the principles of open systems, adhering to recognized standards, using proven solutions, hierarchical organization of IS with a small number of entities at each level – all these contribute to transparency and good manageability of the IS;
2. Continuity of protection in space and time, inability to overcome protection tools, exclusion of a spontaneous or induced transition to an unsafe state – under any circumstances, including abnormal, the protection tool either fully fulfills its functions, or completely blocks access to the system or part of it;
3. Strengthening of the weakest link, minimization of access privileges, separation of functions of maintenance services and responsibilities of personnel [9]. It is assumed that roles and responsibilities are distributed in such a way that one person cannot disrupt a critical process for the organization or create a security breach through ignorance or the order of cybercriminals. With regard to the software and hardware levels, the principle of minimizing privileges implies that users and administrators are only assigned those access rights necessary for them to perform their official duties [10]. This allows one to reduce damage from accidental or deliberate incorrect actions of users and administrators;
4. Separation of defense, a variety of protective equipment, simplicity and controllability of the information system and its security system. The principle of separation of defense requires that we not rely on a single defense line, no matter how reliable it may seem. Physical protection should be followed by software and hardware, identification and authentication by access control, logging and auditing. Layered defense is able not only not to let the intruder through, but also in some cases to identify them by logging and auditing their details [6]. The principle of a variety of protective equipment implies the creation of defensive lines of different nature, so that a potential attacker is required to master various and, if possible, incompatible skills.

A thoughtful and ordered structure of software tools and databases, the topology of internal and external networks directly affects the achieved quality and security of the IS, as well as the complexity of their development. With strict adherence to the rules of structural design, it is much easier to achieve high quality and safety indicators, since the number of possible errors in implementing programs, equipment failures and other failures is reduced, their diagnostics and localization are simplified. In a well-structured system with clearly identified components (client, application server, resource server), the checkpoints are clearly distinguished, which solves the problem of proving the sufficiency of the applied protection means and ensuring the impossibility of circumventing these means by a potential intruder [5].

The following are considered as objects of vulnerability:

- Dynamic computational process of data processing, automated preparation of decisions and development of control actions;
- Object code of programs executed by computational means in the course of IS functioning;
- Data and information accumulated in databases;
- Information provided to consumers and to actuators.

The task is to identify the factors on which the listed threats depend, in the creation of methods and means to reduce their impact on IS security, as well as in the effective allocation of resources to ensure protection of equal strength to all negative influences.

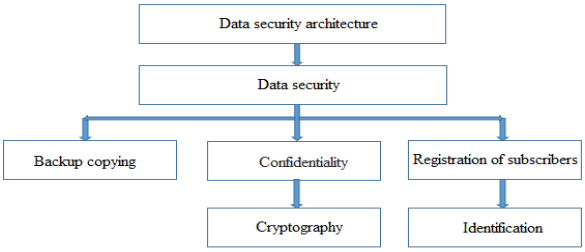

Fig. 3. Data security architecture development

5. Crash tolerance analytics

It should be emphasized that no matter how powerful security systems are, they cannot guarantee complete reliability of the software and hardware level tolerance to malfunctioning. In this section, we focus on the latter: the host crashes due to hardware related causes. We assume that the application server software is reliable and that the DBMS and archival systems are also reliable and do not crash. We focus on how to cope with crashes of front-end application servers that carry out the tasks of clients – both inside and outside the enterprise, by connecting to the DBMS [8].

The rationale for our starting set of reliability assumptions are two-fold: the DBMS is generally procured from high-quality vendors who offer their own reliability and maintenance guarantees and application server programs are typically developed using well-known design methodologies and are tested prior to deployment. What we cannot control are application server hardware faults that can develop as a result of deterioration over time and, when they lead to a host crash, any clients' work in progress is interrupted requiring those clients to repeat the document processing activities [7].

Recall that we already addressed scalability or load balancing concerns by having H (> 1) servers in the front end and thereby limiting the average number of customers per application server host to a small value of N/H. This use of multiple application server hosts will also provide a level of crash tolerance as the customers connected to a crashed host can reconnect to one of the operative hosts and resume their work. However, these reconnecting customers will have to re-execute their entire work using the new hosts, as all the work they did in the crashed server will have been lost. This extra work can be substantial for N/H customers. To see this extent of work loss, let us first model the operational behavior of application servers:

- A client connects to a host and instantiates an application server;
- The server fetches the required documents from the data tier and caches them locally at the host; this involves data transfer from the DBMS to the host;
- The client works on the cached documents for some time;
- The modified document versions are then uploaded to the data tier; this involves data transfer from the host to the DBMS.

Purely to motivate the challenges posed by host crashes, assume that a client's work time is constant and W (>0) time units. Note that if the client works only in read-only mode, no document needs to be uploaded; we ignore these as edge cases for now.

If a host crashes just before a client caches the documents from the DBMS, no work is lost as no work has started; on the other hand, if the host crashes just before a client uploads the modified documents to the DBMS, all the work carried out so far is lost and that client will have to re-start from scratch on the reconnected host. We assume that when a host crashes, the wasted work per client connected to the crashed host is uniformly distributed on (0,*W*); the expected amount of wasted work will therefore be *W*/2 per client. Since the average number of clients connected to a crashed host is *N/H*, the average total amount of wasted work will be (*N/H*)(*W*/2). For example, if we let $N = 20$, $H = 2$, and $W = 10$ minutes, then the average total amount of wasted work that needs to be re-done amounts to 50 minutes. We make the following observations regarding service interruptions when host crashes are allowed:

1. Use of multiple hosts for scalability does not completely address the challenges posed by host crashes;
2. When a host crashes, the load on operational hosts increases: *N/H* increases to *N*/(*H*-1); and,
3. When *W* is large, the amount of lost work to be repeated can be substantially large.

While 1 and 2 are inevitable, the challenge we will next address therefore will be to explore ways of minimizing the amount of lost work due to a host crash.

6. Conclusion

The developed information system is designed for the efficient operation of the administration (user, client) of the enterprise. The client application has the two possible types of connection to the enterprise information system: desktop client and web client. The desktop-client application enables connection to the information system on the working computers of the enterprise through the local network, with an appeal to the application server. Outside the enterprise, the user can work with the information system through web browser, as a web client, with an access to a web server. The developed multilevel architecture, Enterprise digitalization system. Document processing schemes, Document approval process, Data Security Architecture are the guarantor of the availability, reliability and security of the system.

Acknowledgments

This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant #AP09259208)

References

[1] Andersson J.: Enterprise Information Systems Management. Kungl Tekniska Högskolan, Stockholm, 2002 [https://www.diva-portal.org/smash/get/diva2:9165/FULLTEXT01.pdf].
[2] Balakayeva G. T. et al.: Using NoSQL for processing unstructured Big Data. News of the National Academy of sciences of the Republic of Kazakhstan 6(438), 2019, 12–21.
[3] Balakayeva G., Darkenbayev D.: The solution to the problem of processing Big Data using the example of assessing the solvency of borrowers. Journal of Theoretical and Applied Information Technology 98(13), 2020, 2659–2670.
[4] Duffy D.: Domain Architectures. Models and Architectures for UML Applications. Datasim Education BV, Amsterdam, Netherlands, 2004.
[5] Gill M.: Issues for consideration in mergers and takeovers from a regulatory perspective. BIS Review 60, 2000 [https://www.bis.org/review/r000721b.pdf].
[6] https://downloads.cloudsecurityalliance.org/assets/research/security-guidance/security-guidance-v4-FINAL.pdf
[7] https://essay.utwente.nl/59810/1/MA_thesis_J-W_van_Houwelingen.pdf
[8] https://resources.sei.cmu.edu/asset_files/WhitePaper/2006_019_001_52113.pdf
[9] https://www.imageapi.com/blog/digital-document-management/
[10] https://www.kreyonsystems.com/Blog/digitization-with-document-management/
[11] ISO 15704:2000. Industrial automation systems. Requirements for enterprise-reference architectures and methodologies.
[12] ISO 35.100. Open systems interconnection.
[13] ISO/IEC 2382:2015 Information technology, Vocabulary.
[14] ISO/IEC 42010:2011. System and software engineering – Architecture description, 2011.
[15] Katsaros G. et al.: A Multi-level Architecture for Collecting and Managing Monitoring Information in Cloud Environments. 1st International Conference on Cloud Computing and Services Science, Noordwijkerhout, 2011.
[16] Keith Stouffer K. et al.: Guide to Industrial Control Systems (ICS) Security. NIST Special Publication 800-82, 2015 [http://doi.org/10.6028/NIST.SP.800-82r2].
[17] Panarello A. et al.: Blockchain and IoT Integration: A Systematic Survey. Sensors 18, 2018, 2575 [http://doi.org/10.3390/s18082575].
[18] Panetto H. et al.: New Perspectives for the Future Interoperable Enterprise Systems. Computers in Industry 79, 2016, 47–63 [http://doi.org/10.1016/j.compind.2015.08.001].
[19] Sen J.: Security Privacy Issues in Cloud Computing Computing. [https://arxiv.org/ftp/arxiv/papers/1303/1303.4814.pdf].

Prof. Gulnar Balakayeva
e-mail: gulnardtsa@gmail.com

Doctor of Physical and Mathematical Sciences, Professor of the Department of Computer Science, Faculty of Information Technologies, Al-Farabi Kazakh National University, Almaty, Kazakhstan.
Research Interests: Development of big data processing systems, modeling of physical and chemical processes.

http://orcid.org/0000-0001-9440-2171

Prof. Paul Ezhilchelvan
e-mail: paul.ezhilchelvan@ncl.ac.uk

Ph.D., Professor: University of Newcastle upon Tyne, Newcastle, United Kingdom.

http://orcid.org/0000-0002-6190-5685

Ph.D. Yerlan Makashev
e-mail: makashev.yerlan70@gmail.com

Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Computer Science, Faculty of Information Technologies, Al-Farabi Kazakh National University, Almaty, Kazakhstan. Research Interests:Modeling of physical and chemical processes, Development of computing systems.

http://orcid.org/0000-0003-1698-3614

Prof. Christofer Phillips
e-mail: chris.phillips@newcastle.ac.uk

Ph.D., Professor: University of Newcastle upon Tyne, Newcastle, United Kingdom.

http://orcid.org/0000-0002-2470-1659

Ph.D. Dauren Darkenbayev
e-mail: dauren.kadyrovich@gmail.com

Ph.D., Associate Professor of the Department of Computer Science, Faculty of Information Technologies, Al-Farabi Kazakh National University, Almaty, Kazakhstan. Research Interests: Big data processing. Development of computer systems for the educational process.

http://orcid.org/0000-0002-6491-8043

Ph.D. Kalamkas Nurlybayeva
e-mail: kalamkas.nurlybayeva@gmail.com

Faculty of Information Technologies, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

http://orcid.org/0000-0002-2069-2564