

Evolution of legacy systems : strategic and technological issues, based on a case study

Herman Tromp and Ghislain Hoffman
Department of Information Technology
Ghent University, Belgium
Herman.Tromp@UGent.be, Ghislain.Hoffman@UGent.be,

Abstract

The goal of this experience report is to highlight the strategic and managerial issues that are likely to be involved in any migration or evolution effort of large-scale legacy systems. The report is based on a specific case study of a large organisation in the Belgian health care and social security system. Major drivers for a legacy evolution effort are identified. Emphasis is put on the required management and planning view, rather than on the mere technological issues. Their constituent elements are discussed in some detail.

Introduction

In a companion paper [1] some first findings of the ARRIBA research project are described and directions for future research are outlined. While ARRIBA tries to define the more technological aspects of legacy mining, knowledge extraction and revitalisation in a generic and long term way, in this paper we report in more detail on specific experiences of a specific case. In [1] only a summary description of this case is presented, besides some other cases. We will also discuss the strategic and managerial aspects, rather than limiting the scope to pure IT issues.

While ARRIBA focuses on research towards techniques for reverse engineering and architectural knowledge extraction, the case we present here is mainly driven by the need to revitalise an IT infrastructure that is crucial for the organisation and is driven by the need to further evolve a large scale legacy application. Some of the most obvious drivers behind that need will be explained below, but it is important to note that they are widely different in nature. They include technology, economy of scale, strategy and market position.

This experience report is based on a spin-off project of the ARRIBA project. This project is a bilateral cooperation between the Department of Information Technology, Ghent University and the LCM (“Landsbond Christelijke Mutualiteiten”). The latter is the largest independent organisation, which is, as a part of the Belgian Social Security system, responsible for providing the redistribution of health care insurance allowances, both towards individuals and hospitals etc. Besides their legally regulated core mission, they also offer a number of welfare related services to their members. It should be noted that their operational and legal context is typically Belgian, which implies that no COTS software can be found on the international market to support their core business. Even if a standard ERP package is installed, it offers only a partial solution, for example in basic accounting operations. Particular legal requirements exclude the use of standard packages to cover all of them, mainly due to fundamental differences in information models and business processes.

Although the case study is about a large-scale migration effort, the best approach, as will be explained below, appears to be an evolutionary. The managerial problems associated with that approach will be discussed in some more detail further in this paper.

For reasons of confidentiality, some details in this experience report are made somewhat more generic, but the conclusions remain sufficiently based on real experience to be relevant for this workshop.

Drivers for legacy evolution

Several alternative definitions of what exactly a legacy system is can be envisaged (see for example [6] [7]). In those definitions, often the notion of “something valuable” is present, as well as the notion of “old, obsolete”. It is clear that legacy systems are crucial for the operation of organisations, which are essential for our economical and welfare activities. If we want to identify, however, what the drivers and the needs for evolution are, we prefer to use a more pessimistic definition:

A legacy system is an operational system that has been designed, implemented and installed in a radically different environment than that imposed by the current IT strategy.

A careful analysis of the above definition allowed us to identify a number of most evident drivers for having the system migrate and evolve. A number of them are listed below and illustrated by the case.

- The corporate strategy gets redefined, e.g. from a traditional data processing model to a multi-channel, service oriented model. This goes together with the requirement to have up-to-date data, coming from multiple sources, online all the time. It must be noted that currently, data are often replicated (daily, weekly) at local offices, for historical reasons (lack of sufficiently performing communication infrastructure, unclear definition of data ownership, etc.). Historical data are often only available off-line, e.g. archived on tape.
- Legal requirements and regulations in Belgian health care insurance change often, and those changes hardly ever take into account the IT system characteristics. As an example of this, new legislation requires the use of archival data, which is currently only available on magnetic tape, in order to impose a limit on the maximum health care cost per household (a notion which is not strictly defined, by the way), depending on their total income, requiring an interface with the taxation services.
- Business processes are redefined when management and business structure is reorganised
- The total cost of ownership of current systems becomes prohibitive, due to the diversity of the systems and the cost of software maintenance. On top of that, due to a growing business volume and the data processing model used, performance becomes increasingly an issue, raising the question whether to invest either in more powerful, but expensive hardware or to migrate to a new hardware/software platform with a larger evolutionary capacity.
- In the case we were also confronted with an outright end-of-life situation (no proper data base system, phasing out of a line of hardware and system software, etc.), resulting in an absolute need to migrate to a new platform. It should be noted, however, that the timeframe in such a situation is still a few years, but not very much longer
- Obstacles for migration can clearly be identified
 - A corporate information model does not exist. That model is deeply hidden in a proprietary flat file system: essential corporate data are stored in a single file, which is accessed through a wrapper. That wrapper performs maps the logical view on the data to their physical structure, but also accounts low-level tasks such as hashing, garbage collection, memory mapping, etc. The original developer of that “data structure” has retired, which makes it extremely hard to recover the data model.
 - There is no well defined IT architecture

Other factors hindering evolution, as already mentioned in [1], were also found in this case. Just to mention a few:

- The predominance of COBOL code, which has severe implications
 - Knowledge of the code is getting lost (the experts are retiring)
 - Recent techniques for software reengineering [5] are often based on object oriented languages and it is not clear yet how they can be fully used in this environment
- The project driven nature of the development efforts in the organisation often prevents a uniform, organisation-wide view

- Recent technologies, based on standard ERP packages [11], EAI techniques [8] [9] [10] [12], data warehousing, etc are often not very well understood. This is mainly due to a significant gap between the business view on organisation needs on one hand, and the IT infrastructure on the other hand.

A feasible migration strategy and a management view should seriously address the above. In the next section, we will concentrate on those issues.

Management view and plans

In order to support a revitalisation effort in a large organisation, management must clearly formulate answers to the following questions

- What is the motivation for the effort? (Why do we do it?)
- What are the objectives of the migration effort? (Where do we want to get?)
- What are the basic postulates and constraints? (What do the environment and previous management decisions impose?)
- How do we measure success, both on the road and at the end? (What are the critical success factors?)
- What are the risks? How to assess and address those risks?
- What methodology do we adhere to? (How do we get there in a systematic way?)

In the following subsections, we will discuss those issues more specifically for the case at hand.

Motivation

A concise motivation is required to get the stakeholders' and management buy-in. In the case at hand, the following elements are at play:

- A technology drive is present, but should not be overestimated. The main issue with this respect is to make sure that the organisation develops and maintains a *strategic* technology, while keeping aligned with industry development, and employs the selected technologies appropriate to reach its long-term objectives
- The business driven motivation is the aim to remain the national leader in health care and related social services
- The current hardware platform (BS2000) is reaching the end of its useful lifecycle and must be replaced anyhow
- Migration involves some risks, but appears to be essential and inevitable for future evolution
- The strategic decision to start with a "REFAC" project (Reorganisation of the Financial, Administrative and Control Circuits) implies a complete revision of the health care information system and is the basis for a reengineering effort
- Migration to a new environment should reduce support requirements, enable faster response times in development needs and allow to reassign IT staff to support emerging new technologies
- Viewing data and information as an institutional asset will improve the quality of (management) reporting and allow staff to respond easily to rapidly expanding needs for information, as well as provide a service oriented environment.
- In this view, it is felt that a more structured one should replace the present proprietary data infrastructure. A relational data model seems appropriate, but performance remains a major issue and this could impose constraints on the feasibility of deploying a relational database system.

It should be clear that the above list is a mixture of technological, management and strategic issues. As an expression of interest and motivation, they provide a clear and necessary commitment from

management, which, by the way, consists mainly of non-IT professionals but rather of medically trained people.

Objectives

In this section of any master plan for migration and evolution, the desired outcome and objectives must be formulated, based on both the “motivation” section but taking into account the environment, as described in the “postulates and constraints” section.

In the case at hand, the objectives were phrased as follows:

- Set up a migration path from the present operating platform (BS2000) to a new one. This is not only a major technological challenge, but requires also a profound economical analysis.
- Define a new technical architecture to be developed and installed. Several options are open for consideration.
- Create an efficient and flexible application and data architecture, which is felt to be lacking currently
- Accompany this with a budgetary, human resources management and business reengineering plan
- Make the necessary budgets available. This point depends heavily on management and stakeholders’ buy-in, as defined in the “Motivations” section

The above list is definitely not exhaustive, but turned out to be both sufficiently concise and elaborate to convince the stakeholders, including the users, that their interests are best served.

Postulates and constraints

This section lists constraints imposed by the present environment and by previous management decisions. It should be noted that a previous reengineering project has failed, at a considerable expense, so a clear understanding of those constraints must be stated. Some of the previous decisions can definitely be argued, but license costs of software packages already incurred must be given consideration in the final cost/benefit analysis.

- A big bang migration is not feasible, because of ongoing operational requirements and obligations.
- Migration must be gradual. This implies that more is needed than a mere refactoring exercise, since data migration, synchronisation and consistency are major issues
- In every migration step, data must remain synchronised. This involves inevitably some degree of replication in intermediate stages, but finally it should be avoided. A central data provider must be present, instead of replicating data asynchronously to local offices.
- A business process reengineering exercise is going on and it is not very clear how this can be synchronised with IT migration
- A corporate information model must be developed. In [2] is described how such a model can be extracted from existing persistent data structures *and* from COBOL code, but in this case much of the information cannot easily be obtained from COBOL record structures and is embedded in the executable code. The central role of a corporate information model within an organisation is explained in [3].
- This corporate information model should be based on an existing relational database product, which was chosen a few years before, for mainly commercial reasons
- Real-time and on-line data access (24 x 7) becomes a stringent requirement for the future and service oriented environment
- A J2EE based thin client architecture through portal and application servers is deemed to be most appropriate, but few experience (estimated at 5 over 200 traditional COBOL developers) is available. This turns out to be a major challenge.

- Integration with newly developed customer relationship management and financial systems (already based on newer technologies) is required

All these put severe constraints on the feasibility of possible migration paths.

Critical success factors (CSF)

The success of an evolution effort has to be measurable. Therefore, a number of critical success factors (CSF's) must be defined. Those should not only be used to evaluate the final outcome, but should be used along the road to measure progress and to define decision points in the plan, where go/no-go decisions should be made.

In this particular case, the following CSF's were identified:

- An affordable but significant proof of concept must be delivered within a reasonable amount of time and effort (typically 4 months)
- Along the whole roadmap, quick wins must be identified to validate the migrating system, and those should offer a measurable business benefit
- A strategic implementation of a strategic application suite within a time-frame of two years is essential
- Support from the stakeholders must continuously be ensured
 - From management
 - From users, both corporate and individual

This list could easily be enlarged, but on the other hand reflects what was mentioned in the "Motivations" section.

Risk assessment

A number of risks are definitely present in a large-scale migration plan, and must clearly be identified in order to be controllable. A list of risks is indispensable in any plan, but is certainly preliminary. New risks are likely to pop up.

Possible risks in this case are:

- Unfeasibility. Before a proof of concept or even a pilot application is delivered, it remains uncertain whether the combination and integration of traditional and "new" technology is feasible. The difficulties involved in controlling transactional context in a mixed environment are, for example, discussed in [4].
- Especially the interaction between a COBOL environment and a Java environment is a major technological risk factor. See also [4].
- Complexity. Systems may tend to be overly complex, and complexity is a source of unmaintainability. Integration of systems of a different nature does not relieve this risk. Componentisation and loose coupling between constituents might bring a solution, but well-understood solutions are not available yet.
- Performance is already a problem in the "as-is" situation, mainly because of the data processing mode used. To control costs and guarantee operational flexibility, it remains a major issue in the new environment, especially while transitioning.
- Obsolescence of techniques. Even "new" technology tends to be obsolescent before it even matures.
- Lack of pragmatism, taking into account the current situation in terms of expertise, resources and technology available. Solutions have to be pragmatic. Solutions offered by academics often fail in this respect – we must admit that. We should clearly attend to what was presented before in the "Postulates and constraints" section.

- Unmanageable systems. This point is related to “complexity” as mentioned before. What is overly complex is also unmanageable. There is also a human resource issue here: do we have the people and staff to manage those environments?
- Cost control. This point might seem to be obvious. but cost estimates in legacy evolution are hard to obtain.
- Lack of appropriate (human) resources, both within the organisation as on the local market. This point was mentioned earlier. It is clear that finding the right people and skills is a major impediment.

In the present case, the risks were carefully evaluated and fallback positions were defined.

Methodology

A methodological action plan is required.

In the case at hand, a number of tracks were defined to plan the required actions. These are

1. Evaluation of the “as-is” situation in the current mainframe environment. Relevant action items are among others: evaluation of the existing COBOL code and underlying data structures
2. Determine the “to-be” business and application architecture
3. Determine the “to-be” technical and deployment architecture
4. Critical evaluation of the migration scenario’s between the “as-is” and “to-be” situations

The careful development of this methodology allowed (and still is allowing) the streamlining of the migration path. Although the division between the tracks seems, at least at first sight, to be rather artificial, it partially reflects the structure of the different teams in place and so it is a consequence of the organisational structures, which cannot be ignored.

Technological migration issues and solutions

Based on the above managerial and tactical considerations, several technical migration scenarios are examined at the moment. For company-confidential reasons, those cannot fully be exposed right now, but that can very soon be remedied and result in a more expanded version of this position paper. underpinned by facts and figures.

Possible solutions to the evolution and migration plan will definitely have to rely on a temporary and controlled form of mirroring of data between the current mainframe platform and the relational database to be deployed.

Conclusion

Any effort for migration or evolution of large-scale industrial software systems must be driven by a corporate strategy redefinition. Buy-in from management and end users requires a well-defined strategic management view and planning, based on a clear statement of motivation, objectives and of the constraints imposed by the current environment and continuous operational needs. The definition of measurable success factors and a precise risk assessment are also essential. An organization-wide methodology must be defined and supported by all stakeholders. As a final note, it should be pointed out that the case on which this paper is based is a not- profit organisation. It is likely that in a more commercial context, a more detailed cost-benefit analysis would be needed.

References

- [1] I. Michiels, D. Deridder, H. Tromp, A. Zaidman, “Identifying ICT problems in legacy software: preliminary findings of the ARRIBA project”, this workshop.
- [2] J. Henrard, J-M. Hick, P. Thiran, J-L. Hainaut, “Strategies for data reengineering”, Proc. WCRE02, IEEE Computer Society Press, 2002
- [3] K.Vandenborre, P. Heinckiens, G. Hoffman, H. Tromp, “Coherent Enterprise Modelling in Practice”, 13th European-Japanese Conference in information modelling and knowledge bases”, Kitakyushu, Japan, 2003.
- [4] D. Plakosh, S. Comella-Dorda, G.A. Lewis, P.R.H. Place, R.C.Seacord, “Maintaining transactional context: a model problem”, Report CMU/SEI-2001-TR-012, Carnegie Mellon Software Engineering Institute, August 2001
- [5] S. Ducasse, S.Demeyer and O.Nierstrasz, “Object-Oriented Reengineering Patterns”, Morgan Kaufmann and Dpunkt, 2002.
- [6] M.L. Brodie and M. Stonebraker, “Migrating Legacy Systems – Gateways, Interfaces and the Incremental Approach”, Morgan Kaufmann Publishers, 1995.
- [7] Aberdeen Group, “Legacy Applications: from cost management to transformation”, Executive White Paper, March 2003, at <http://www.aberdeen.com/2001/research:03038126.asp>
- [8] D.S.Linthicum, “Enterprise Application Integration”, Addison-Wesley, 1999
- [9] J.C.Lutz, “EAI Architecture patterns”, EAI Journal, March 2000
- [10] M. Themistocleous and Z. Irani, “Evaluating and adopting application integration: the case of a multinational petroleum company”, Proc. 35th Hawaii International conference on system sciences, 2002.
- [11] M. Themistocleous, Z. Irani, R.M. O’Keefe and R. Paul, “ERP problems and application integration: an empirical survey”, Proc. 34th Hawaii International Conference on system sciences, 2001.
- [12] M. Fowler, “Patterns of Enterprise Application Integration”, Addison-Wesley Signature Series, 2002.