SIM: a working example of an e-government service infrastructure for mountain communities

(Extended Abstract)

Franco Arcieri[1] Elettra Cappadozzi[2] Enrico Nardelli[3] Maurizio Talamo[2,4]

1. Consultant to AIPA for the SIM ("Sistema Informativo della Montagna") project.
2. "Coordinamento dei Progetti Intersettoriali" of AIPA - "Autorità per l'Informatica nella Pubblica Amministrazione", Roma, Italia.
3. Dipartimento di Matematica Pura ed Applicata, Univ. of L'Aquila, L'Aquila, Italia. & Istituto di Analisi dei Sistemi ed Informatica, C.N.R., Roma, Italia. nardelli@univaq.it
4. Dipartimento di Matematica, Univ. of Roma 2 "Tor Vergata", Roma, Italia. talamo@aipa.it
M.Talamo is the CEO of the Initiative for the Development of an IT Infrastructure for Inter-Organization Cooperation (namely, "Coordinamento dei Progetti Intersettoriali") of AIPA - "Autorità per l'Informatica nella Pubblica Amministrazione".

Abstract

In this paper we present an Information Technology (IT) based infrastructure providing e-government services to people living in mountain areas. Such a system, SIM - "Sistema Informativo della Montagna", provides to citizens in these areas services in the cadastral sector, in the labour and pension sectors, and services related to public registries for personal data. SIM is an example of a general distributed IT infrastructure, introduced in [1]*, able to integrate and to make heterogeneous information systems interoperable, while providing good performance levels. SIM is currently being used in about one thousand operating centers, serving ten millions inhabitants and about half of the italian municipalities and it is currently being designed its extension to the whole italian territory.*

Design and development of SIM have been conducted by the italian Ministry for Agriculture and Forest Policies within an overall initiative, coordinated by the "Coordinamento dei Progetti Intersettoriali" of AIPA, the Italian Authority for Information Technology in Public Administration, aiming at developing inter-organization cooperative information systems supporting e-government actions.

Keywords: federated information systems, information systems interoperability, e-government information technology infrastructure.

1 Introduction

A critical issue for the realization of every e-government action is that usually services to be provided to end-users require the interaction between IT systems developed over the years by autonomous and independent organizations. This interoperability problem is at the heart of the effectiveness of every effort in this field. Over the last couple of years, given the explosion of Internet connections, the importance of such a scenario has reached a peak and the topic of integration of Information Systems has become a really hot one [9].

An important characteristic of this scenario is that the involved organizations share a common semantic background and are available to make some of their resources available to support cooperation, but cannot engage themselves in a full re-engineering or re-implementation effort. The challenge is therefore how to keep the *coherence of the overall (distributed) set of data* during interaction of legacy systems. On one side, in fact, data are independently and autonomously managed by the various organizations. On the other one, data are needed and used also outside the organization producing/managing them and controlling their changes. These clashing situations will produce incoherence in the overall set of data, sooner or later, with absolute certainty.

Since the lack of coherence derives mainly from the organizational framework, then the technical solution has to be designed in a way to match needs and behaviour of the organizations involved. Moreover, the technical solution for coherence maintenance has to be designed so that the overall system has good performances and both technical and organizational costs of cooperation are not hidden.

A recent series of workshops on engineering federated database systems and information systems [6, 7, 8] has pointed out that one of the open research issues in this area is the integration of legacy databases in a federation of autonomous heterogeneous information systems and the ability of maintaining coherence of information representation during evolution over time of the systems involved in the federation.

A widely followed approach to issues regarding data integration in multidatabase system is based on the *wrapper-mediator* architecture [12] coupled with an object-oriented approach (e.g., [5]). But the main drawback of such an approach in large-scale and legacy systems like the ones found in the scenario above described is that to consider the existing systems as black boxes may be catastrophic in terms of performances. In such a case, in fact, given the wrapping provided by the OO technology, access number and access paths required to an underlying Source Database by the execution of coherence maintenance functions are largely out of the control of the designer. Thus, providing acceptable performances is a highly challenging task [10].

Our approach, on the contrary, makes it possible to evaluate and tune the impact on performances of a given Source Database deriving from outside requests. Hence our approach makes it possible to perform a rightsizing of the overall system through a cost-benefit analysis.

We have addressed this interoperability issue by defining a a general model (the **Access Keys Warehouse** approach) focused on direct inter-organization cooperation and have developed an new architectural approach aiming at this target [1]. Examples of large-scale systems designed according to this approach are discussed in [2, 3, 4].

Our approach makes interoperability costs visible and supports their negotiation in a flexible way. Hence its major organizational impact is to bring to the surface the hidden costs of inter-organization cooperation.

SIM is an example of an implementation of the Access Keys Warehouse approach focused on supporting decentralization policy in Public Administrations services. It operates in mountain communities and acts as a mediator between citizens living in those areas and IT-based services in the fields of cadaster, labour and pension, public registry of personal data.

In this paper we describe its logical and functional architecture and services it provides. Design of SIM started in 1998 and the overall financial effort has been, until now, of about 100 billion liras (roughly 50 million US dollares). It is currently being used in about one thousand operating centers, all over Italy, serving more than 10 million inhabitants and more than 4000 of the about 8000 municipalities. Its extension to the whole country is currently being planned.

2 Architecture

System architecture of SIM is a physically distributed one. This is compliant with more recent trends and requirements in Public Administrations and e-government actions where decision capabilities are increasingly and increasingly de-centralized and put at the appropriate local level.

Hence SIM is structured around Service Centers, providing services to end users, coordinated by various Regional Centers. A national Managing Agency is then in charge of the overall control, coordination and certification activities.

From a functional viewpoint a SIM Service Center is composed by the following subsystems (see also figure 1):

- external database access,
- workflow management,
- legacy system access,
- certification and security,
- information flows correlation.

We now describe their functions in more detail:

Access to external databases. This subsystems manages all information flows towards databases external to SIM but needed for a correct execution of its services. Hence this subsystems manages and executes procedures to connect to and to interact with external databases: this the unique access point to such external information sources for all SIM subsystems.

All technical constraint are encapsulated and resolved within functions of this subsystem. E.g., the need to use specific terminal emulators (3270, vt100, ...), requirements on access and communication protocols (X25, ISDN, SNA, ...).

Management of workflow procedures. Support of organizational procedures related to specific services is provided by this subsystem, which manages and coordinates information flows needed to supply SIM services. For example, in the case of a building permit request by a citizen, this subsystem manages requests to external organizations - when their opinion is required by the law for such a permit - and their answers, and keeps track of the state of the request.

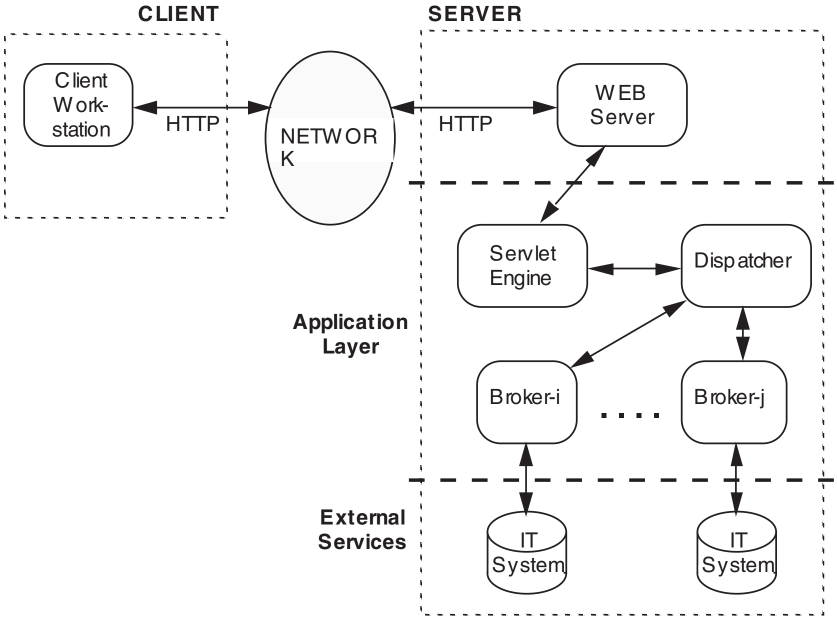

Figure 1. Overall functional architecture of a SIM Service Center

Moreover, this subsystem coordinates all information flows internal to SIM, flows that starting from any access point may, in principle, affect every SIM subsystem.

Access to legacy information systems. Querying information systems of Public Administration is usually restricted both functionally and technically. In fact, queries have a limited number of different types and structures (e.g., very few partial match requests can be satisfied) and have organizational constraints (e.g., each access has a cost depending on the use of the retrieved information). This kind of problem is encapsulated within procedures of this subsystem.

Certification and security procedures. This subsystem is in charge of security levels of SIM: in fact, it provides to other subsystems that require them, all security and certification services needed to correctly supply to end-users SIM services. It also automatically provides documentation about internal information flows supporting workflow and legacy subsystems.

Correlation of information flows. Each information flow traversing SIM brings with itself information which is potentially correlated with other pieces of information known to SIM but is not directly related to the purpose activating the specific information flow. Since SIM is a mediator between different services is important for it to establish and to check correct correlations between all information flows traversing it and information sources known to it. This task is managed by this subsystem.

3 Certification and Monitoring

Certification is a critical activity for any IT-system supporting e-government. Given the official value of data and the legal role they play in the Public Administration context, it is mandatory to be able to prove transmittal and receival of data.

SIM therefore features sophisticate functionalities for certification, documentation and accounting of information flows between clients and servers (see also figure 2). These are based on *network probes*, operating both on server side and on client side, and controlling all traffic flowing between them and the communication network.

Network probes are devices featuring high security levels, such as absence of terminal devices to access internal resources (e.g., keyboard, mouse, screen, ...), uninterruptable power supply, software components for automatic faults check, diagnosis and alert. Their action is very efficient and does not slow down operations of the communication layers.

Probes are able to detect all data packets related to application services passing on the portion of the network they are controlling and to select only those referring to specific kinds of transactions. From these data packets, probes then extract and store suitable data item able to certify and to document information transmitted and received at both sides of a communication flow.

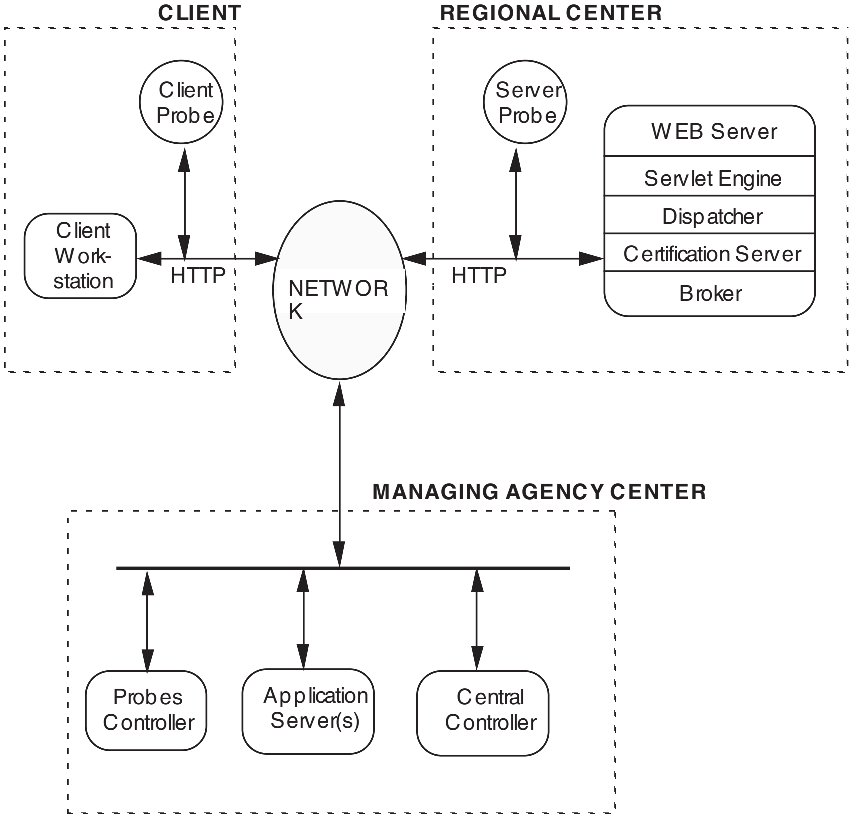

Figure 2. Architecture for certification services

Data packets extracted by network probes are sent to a central *probes controller*, which processes all received items and establishes correlation among those referring to the same information flow. These correlations are also stored in a relational DBMS for reporting and statistics purposes. In such a way a complete documentation for each service is built and can serve as the official basis for accounting of services. For some of them, in fact, a fee is asked from the provider and there is the need of certifying the actual supply of the requested service and its correct receival by the client. This certification action is carried out by a *certification server* on the basis of correlation among exchanged data items established by the probes controller.

As said above, documentation about information flows is useful not only for certification purposes: it is critical for performance tuning and monitoring actions and for establishing and controlling the actual levels for quality of services.

4 Information flows

Subsystems that are more critical for SIM for what regards the overall techno-organizational framework are the subsytem for the management of workflow procedures and the one controlling certification and security procedures. These two subsystems, in fact, play an important role in all services provided to SIM end-users. They have to check every information flow traversing SIM to:

- identify user, its work-station and its access privileges with respect to requested services,
- decide and activate security levels adequate for the requested service,
- check correctness and completeness of service requests,
- identify which SIM subsystems are affected by the request and forward requests towards identified subsystems,
- find which external organization is responsible for services provided by SIM in account of them,
- document state of information flows and progress of requests.

To allow a better understanding of architecture and information flows we now present an example of a service request. An end-user connects to SIM web server (running on an NT server) through its web client (Internet Explorer or Netscape Navigator), asks for a specific service and sends needed parameters. Client request is forwarded to an application server (Servlet or Jsp) which is executed within a *servlet engine*. This is done by means of a specific control component, called *dispatcher*, having the task of checking authorization levels of clients against rules defined by the Regional Center. This allows to enable client to consume only those processing resources, both local and remote, for which it has been explicitly authorized. The application server, once received parameter values for service, forwards the request - passing once again through the dispatcher to check authorization levels - to the interested *service provider* in an external IT system, activating it through a *broker*. All systems in the application layer are based on standard middleware software components.

Brokers are therefore fundamental components of the overall architecture, since they are "mediators" between software components providing processing services (i.e., external IT systems) and software components using these services (i.e., application servers). In such a way a full independence is obtained between the needed processing task and the logical organization of data or the algorithmic aspects of their processing. An application server needs only to know which broker is able to provide it a specific service, without bothering about details (e.g., communication protocols, organization of data, ...) of how this service has to be requested.

An additional service provided by an existing or a new external IT-system is added to SIM by simply registering it at a broker, together with necessary information needed to activate it.

Information flows are defined so that each of them is fully self-identifiable, that is each SIM subsystem is able to identify its characteristics without resorting to external resources. This requirement therefore constitutes a first security level, since it allows to raise an alarm whenever a non-compliant flow is detected within SIM. The importance of this derives from the fact that the interconnection network between clients and service centers is not a fully private Intranet, but uses for large parts public Internet. The internal structure of service flows is not visible to SIM end-users and is composed by two main parts:

- *application header*: with a constant length, allows flow identification and is present in every flow between a client and a service provider;
- *service parameter*: with a variable length, contains data needed for a correct activation of the required service.

The servlet engine builds a proper application header for requests just received from a client before forwarding them to the dispatcher. The dispatcher subsystem always examines application header to check authorization levels and to verify correctness and completeness of data in the header: in such a way the end-user can be advised, in case of missing or wrong parameters before forwarding its request to the supplier, with an improvement in overall quality of service.

Network probes also use application headers to detect flows they need to monitor and to extract information needed for certification and documentation purposes.

References

[1] F.Arcieri, E.Cappadozzi, P.Naggar, E.Nardelli, M.Talamo: Access Key Warehouse: a new approach to the development of cooperative information systems, *4th Int. Conf. on Cooperative Information Systems (CoopIS'99)*, Edinburgh, Scotland, U.K., pp.46–56, Sep.99, accepted in the *Int. J. of Cooperative Information Systems*, 2001.

[2] F.Arcieri, C.Cammino, E.Nardelli, M.Talamo, A.Venza: The Italian Cadastral Information System: a Real-Life Spatio-Temporal DBMS, *Workshop on Spatio-Temporal Database Management (STDBM'99)*, Edinburgh, Scotland, U.K., Sep.99, Lecture Notes in Computer Science vol. 1678, pp.79–99, Springer-Verlag.

[3] F.Arcieri, E.Cappadozzi, E.Nardelli, M.Talamo: Geographical information systems interoperability through distributed data exchange, *1st International Workshop on Databases, Documents, and Information Fusion (DBFusion'01)*, Magdeburg, Germany, May 01.

[4] F.Arcieri, E.Cappadozzi, E.Nardelli, M.Talamo: Distributed territorial data management and exchange for public organizations, *3rd International Workshop on Advanced Issues of E-Commerce and Web-Based Information Systems (WECWIS'01)*, San Jose, Ca., USA, Jun.01, IEEE Computer Society Press, 2001.

[5] O.Burkhes, A.Elmagarmid (eds.): *Object Oriented Multidatabase Systems*, Prentice Hall, 1996.

[6] S.Conrad, B.Eaglestone, W.Hasselbring, M.Roantree, F.Saltor, M.Schönhoff, M.Strässler, M.W.W.Vermeer: Research Issues in Federated Database Systems: Report of EFDBS'97 Workshop, *SIGMOD Record*, 26(4):54–56, 1997.

[7] S.Conrad, W.Hasselbring, U.Hohenstein, R.-D.Kutsche, M.Roantree, G.Saake, F.Saltor: Engineering Federated Information Systems: Report of EFIS'99 Workshop, *SIGMOD Record*, 28(3):9–11, 1999.

[8] W.Hasselbring, W.-J. van den Heuvel, G.J.Houben, R.-D.Kutsche, B.Rieger, M.Roantree, K.Subieta: Research and Practice in Federated Information Systems: Report of the EFIS'2000 International Workshop, *SIGMOD Record*, 29(4):16–18, 2000.

[9] W.Hasselbring: Information System Integration: introduction to the special section, *Comm. of the ACM*, 43(6):33–38, June 2000.

[10] V.Josifovski, T.Risch: Integrating Heterogeneous Overlapping Databases Through Object-Oriented Transformations, *(VLDB'99)*, Edinburgh, 1999.

[11] F.Arcieri, E.Cappadozzi, G.Melideo, P.Naggar, E.Nardelli, M.Talamo: A formal model for dealing with data interoperability issue, Tech.Rep. 5/01, U.of L'Aquila, Feb.01, submitted.

[12] G.Wiederhold: Mediators in the architecture of future information systems, *IEEE Computer* , 1992.

1529-4188/01 $10.00 © 2001 **IEEE**