Experiences and issues in the realization of e-government services

(Extended Abstract)

F.Arcieri[1] G.Melideo[2] E.Nardelli[2] M.Talamo[1]

1. Consultant to AIPA - "Autorità per l'Informatica nella Pubblica Amministrazione" & Univ. of Roma 2 "Tor Vergata", Roma, Italia. talamo@aipa.it
2. Dipartimento di Informatica, Univ. of L'Aquila, L'Aquila, Italia, & Istituto di Analisi dei Sistemi ed Informatica, C.N.R., Roma, Italia. nardelli@univaq.it

Abstract

Certification of the executed service is a critical issue for an e-government infrastructure. In fact, given the legal value that is often attached to data managed and exchanged by public administrations, being able to document the actual execution of e-services is of the utmost importance. This is made more complex in cases, like it often happens in the public administration sector, where e-services are based on legacy systems managed by autonomous and independent organizations. In this paper we discuss, starting from real-life e-government systems developed in Italy, the introduction, within the standard three tiers architecture for e-services, of new control components, based on efficient algorithmic techniques, providing solutions for this issue.

Keywords: digital government support, legacy application integration, interoperability support.

1 Introduction

The pervasive spreading of the WWW has created a tremendous opportunity for providing services over Internet. The whole area of e-services is becoming a very hot research field (e.g., see [9] or the special issue [17]), with a wide technological and economical impact. In this context, the digital government sector is quickly gaining importance, given its "potential to profoundly transform citizens' conceptions of civil and political interactions with their governments" [10].

E-government services are defined and developed in a complex architectural and technological scenario. In fact, the various public administrations and agencies that are involved are usually autonomously and independently managed. Therefore, even if they have to cooperate to reach a common overall goal, they have developed and run their own computer-based information systems in a completely independent way.

Supporting intragovernmental processes is therefore a main critical component of any Information Technology (IT) infrastructure for digital government and certification is a very important function of this component.

In this paper we first recall the development of real-life e-government services where we have addressed the certification issue and then present and discuss our solutions, which have led to a successfull realization of the e-services. These systems have all been developed within an overall initiative, coordinated by AIPA, the Italian Authority for Information Technology in Public Administration (PA), aiming at developing inter-organization cooperative information systems supporting e-government actions [13].

It is not the focus of this paper a discussion on security issues, which are another very important aspect of digital government. For a detailed discussion of, e.g., authorization and access control functions in the context of an IT infrastructure for e-government see [11].

The paper is structured as follows. Firstly, in section 2, we present context and motivations at the root of our research and application areas where we have tested and refined our solutions. Then, section 3 discusses the reference architecture we adopt in our discussion. Next, in section 4, we present the fundamental technique supporting certification in intragovernmental processes supporting digital government. Finally, section 5 concludes the paper.

Proceedings of the 12th Int'l Wrkshp on Research Issues in Data Engineering: Engineering e-Commerce/ e-Business Systems (RIDE'02)
1066-1395/02 $17.00 © 2002 IEEE

2 Context and Applications

Origin of our research is in legislation [12] entered in force in Italy in the years 1993-94 requiring to set up an information system for cadastral data exchange by means of telematics communication among Ministry of Finance, Municipalities, and Notaries.

At that time computer based systems were already used, by the above organizations, to manage cadastral data and procedures. Hence each of these organizations was already able to provide (a rough form of) its own e-service to end-users.

The overall challenge was therefore to design the architecture of an IT-based system allowing: (*i*) to control and certify the exchange of information between independent PA organizations, each with its own hardware and software systems, and its different organizational procedures, and (*ii*) to ensure that exchanged data was kept coherent in the different PA organizations, even under updates.

No technical solution was available at that time, since either they required to have homogeneous hw/sw systems, or they put one organization in control of the other ones, or both. Note that certificates related to ownership, location, geometry and value of real estates, issued through these IT-based systems, are mandatory in estate's selling transactions and their issue is subject to a fee, paid for by the buyer.

We investigated the situation and were able to define, through a series of prototypes developed during 1995-97, an architectural solution [1] which fully satisfied requirements and needs of all organizations involved. SICC ("Sistema di Interscambio Catasto Comuni"), is a system which allows to exchange cadastral data among the principal entities interested in Italy to the treatment of cadastral information, that are Ministry of Finance, Municipalities, Notaries, and Certified Land Surveyors [16, 2].

The system is accessible nation-wide through a WEB-based interface since September 1998. The effectiveness of its use is demonstrated by the number of administrative transactions successfully completed through the system in year 2001: they are 800.000 per month, corresponding to 60% of the overall transactions related to cadastral services accomplished every year in Italy.

Afterwards we defined SCT ("Sistema di Comunicazione Dati Territoriali"), that is a prototype developed to investigate technical and organizational issues to allow lowering of costs related to geographical information systems development and maintenance, to enable a true reuse of geographical data by many users and many organizations and thus foster the development of a market for territorial data [3, 4].

The design and prototyping of SCT system was carried out by a consortium led by Telecom Italia and comprising three more large companies (namely, ESRI, Finsiel and IBM). Such a project is one of the steps taken by AIPA in the joint cooperation framework with the Region Association and the Municipality Association for coordinating efforts aimed at supporting the growth of the sector of geographical information systems.

The third and most important example is SIM ("Sistema Informativo della Montagna"), that is a distributed systems providing e-government services to people living in mountain areas [5].

SIM 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. Its design started in 1998 and the overall financial effort has been, until now, of about 100 billion liras (roughly 50 million US dollars). It is currently being used in about one thousand operating centers, all over Italy, serving more than 10 million inhabitants, more than 4000 of the about 8000 municipalities and covering more than 54% of the italian territory. Its extension to the whole country is currently being planned.

Finally we would like to notice that the development of two newest italian e-government services, for which the fourth author is the overall supervisor, coordinated by the Ministry of Internal Affairs, namely the *Electronic Identity Card* and the *Integration of Municipal Public Registries of Personal Data*, is being driven according to the solutions described in this and related papers.

3 Reference Architecture

The reference architecture for our discussion of certification issues supporting intragovernmental processes for digital government is clearly a distributed one (see figure 1). This is compliant with more recent trends and requirements in Public Administrations and e-government actions where decision capabilities are increasingly and increasingly decentralized and put at the appropriate local level. A number of *external services* are attached to the Public Administration intranet but are not directly accessible to clients. For this purpose a client has to access the web-site of the *e-service provider* through which external services are then made available.

In a typical example of a service request an end-user connects to the web server of an e-service provider through a web client, asks for a specific service and sends needed parameters.

Within the provider's site, structured according to the usual three tiers architecture, client's request is forwarded to an application server, executed within a *servlet engine*.

The application server, once received parameter values

Proceedings of the 12th Int'l Wrkshp on Research Issues in Data Engineering: Engineering e-Commerce/ e-Business Systems (RIDE'02)
1066-1395/02 $17.00 © 2002 IEEE

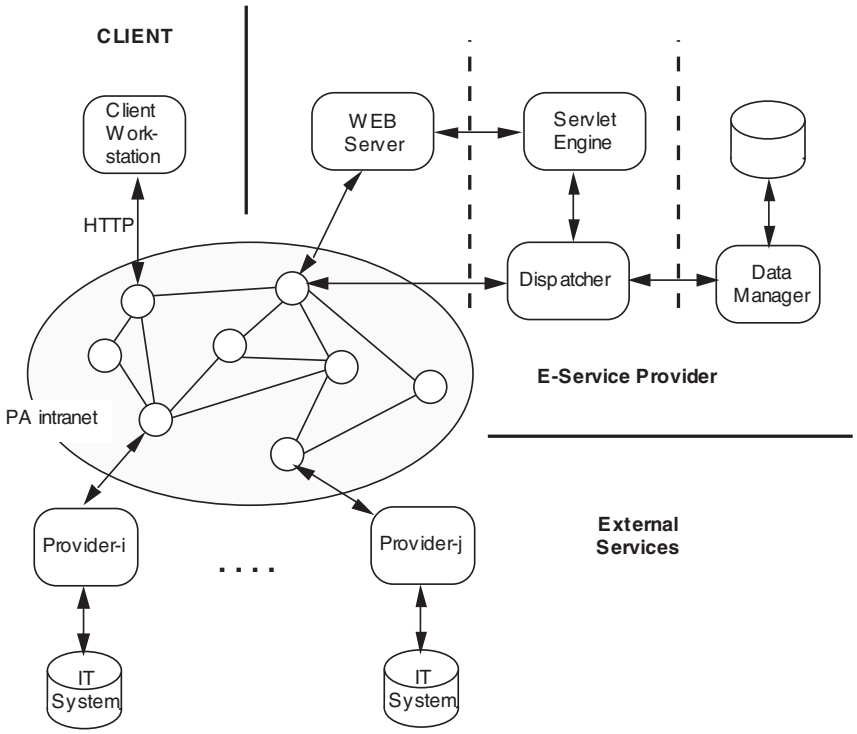

Figure 1. Overall reference architecture

for service, forwards the request to a new architectural component we have introduced, called *dispatcher*. This is a specific control component, having the task of checking authorization levels of clients against rules defined by the e-service provider. More specifically the dispatcher:

- 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 basic e-service providers are affected by the end-user's request and forward them proper sub-requests.

Interaction, for certification purposes, between servlet engine and dispatcher and among all e-service providers is made possible by prefixing exchanged data flows with a suitable *applicative header*, composed by two main parts:

- *flow identifier*: with a constant length, allows to identify characteristics of a flow without resorting to external resources; it 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 structure of applicative header depends on the kind of intragovernmental process which is being supported and it is defined during its design phase. In any case it has to contain: identifiers for the end-user and the workstation he/she is working on, time stamp, transmission parameters, service and function requested. The presence of the applicative header constitutes a first security level, since it allows to raise an alarm whenever a non-compliant flow is detected.

Note that our approach can be integrated with current proposals under development for uniform invocation of remote service, e.g., SOAP [15] or WSDL [18]. In fact the applicative header is simply a format, agreed within the organizational structure of the PA intranet, supporting certification of exchanged services. It can be incapsulated within SOAP or WSDL invocation protocol, if these become standard. Also, the approach is conceptually valid for supporting certification services even over a public network like Internet. Clearly, the role played in our current approach by the applicative header will have then to be implemented through the primitives of protocol(s) that will have been chosen for this purpose.

Proceedings of the 12th Int'l Wrkshp on Research Issues in Data Engineering: Engineering e-Commerce/ e-Business Systems (RIDE'02)
1066-1395/02 $17.00 © 2002 **IEEE**

The servlet engine builds a proper applicative header for requests just received from an end-user before forwarding them to the dispatcher. The dispatcher subsystem always examines applicative 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.

Dispatcher analyzes application threads by means of the same algorithmic techniques applied by network probes (see next section), and is able to quickly check whether an end-user is requiring a service for which he/she has the proper authorization. The introduction of the dispatcher has the great benefit of allowing to deal with client's authorization and authentication issues outside the e-service application programs. The separation among these aspects of e-services means greater flexibility and independence in changing the application logic.

We stress the fact that for an IT infrastructure supporting e-service certification to work efficiently, basic functions have to be independent and abstracted from the actual e-service invoked or executed. Also note that in our approach no wrapping of services takes place but only a monitoring of interaction between client and provider for the aim of documenting what actually happened.

Once the dispatcher has positively executed these checks the request is forwarded either to the third, lowest, tier of the architecture (if only internal data management services are needed), or to an external e-service provider (if another e-service is needed to satisfy the end-user's request).

4 Certification

The main technical difficulty for certification purposes is that all e-services involved are, from the viewpoint of the certification process, like black boxes and cannot be internally changed.

The only approach is therefore to monitor and to keep track of input and output flows. But input and output flows of IT services taken as black boxes are nothing more than sequence of IP packets.

Then the challenge is to be able to understand, by only checking packets flowing through various points over the intranet, what is going on with respect to a specific service. Note that even if we are dealing, in these digital government cases, with the PA intranet, in general all sort of packets can be found. So the goal is to be able:

- at specific points to collect all packets corresponding to specific application and information flows related to a specific service,
- to reverse engineer streams of packets so to reconstruct application level messages passed through monitored points,
- to correlate what has been identified at those specific points so that it can be understood which is the status of the overall distributed transaction,

This approach clearly requires algorithms for packet classification with very high efficiency, otherwise introduces a too large slowdown in the interaction among the basic components. We have a very good algorithm solving this problem [14].

Our architectural solution for certification, documentation and accounting of information flows between clients and servers is based on *network probes*[1] (see also figure 2). They operate both on server side and on client side and, once properly set-up, monitor all and only that traffic flowing between them and the communication network for which they have been explicitly configured. In figure 2 network probes are shown at the site of the e-service provider (server probe) and at the site of its end-user (client probe). Note that, of course, configuration of network probes directly derives from agreements formally established between organizations involved in the exchange of services.

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 neither disturb nor slow down operations of the communication layers. They are physically placed on network links directly leading to or coming from client and server nodes.

Probes are able to detect all data packets passing on the portion of the network they are controlling, to efficiently identify all those related to application services for which they have been configured, and to select only those referring to specific kinds of transactions. From these data packets, probes in fact reconstruct the individual messages of the application threads relative to services they have to monitor, then they extract suitable data items able to certify and to document information passed through the link under their surveillance.

Note that probes neither have to collect and process all data passing on the network link they are monitoring, nor have to reconstruct exchanged flows for all servers on the network. They are configured for specific services and process just IP packets referring to them. Hence this approach is fully scalable without performance degradation. Also note that probes allow a much more efficient support for

[1] Patented, 1997.

Proceedings of the 12th Int'l Wrkshp on Research Issues in Data Engineering: Engineering e-Commerce/ e-Business Systems (RIDE'02)
1066-1395/02 $17.00 © 2002 IEEE

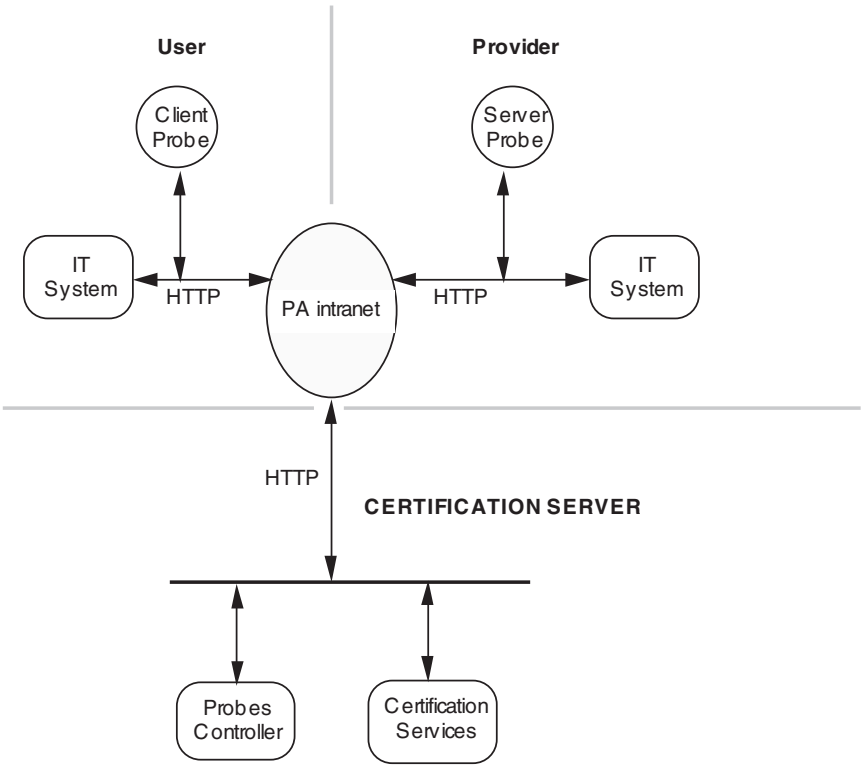

Figure 2. Architectural solution for certification

certification than logging every exchange flow at the dispatcher. Beyond the computational overhead this choice would generate, there would also be the additional complication of having each exchange flow related to the execution of an e-service to pass through the dispatcher, even if it could go directly from an external provider to the end-user.

Synthetic information items filtered out by network probes as result of their real-time analysis of IP packets according to their configuration parameters are sent to a central *probes controller*. This processes all received information items and establishes correlation among those referring to the same service flow. Note that with this approach network traffic overhead is very low since only synthetic elements, at the application level of the communication, and only those relative to the service for which the probe has been configured, are sent back to the probes controller.

The probes controller is then able to aggregate information items referring to the execution of a given service. Probes also compute, on the payload part of packets, a suitably defined hash function, whose outcome is sent, together with other information, to the probes controller. In such a way the controller is able to check also the correct exchange of payload data between provider and end-users, hence to provide a certification of the integrity of exchanged data. These correlations are also stored in a relational DBMS for reporting and statistics purposes.

With this solution a complete documentation for each service is built and can serve as the official basis for certification of services. Whenever a legal value is associated to data there is in fact 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.

An important aspect of this solution based on network probes is that it allows certification and documentation services to be allocated, if desired, to a third-party, independent from both the servers and the clients involved in the exchange of e-services. Note that this does not mean to send externally all IP packets of PA since the third party is anyhow within the PA's intranet. But with this choice the technical leverage point to enact, control, and enforce regulation policies for e-government services is obtained. 'If', 'when', and 'what' are of course issues of competence of higher-level policy makers, but it is important to stress that this technical solution is transparent with respect to the in-

Proceedings of the 12th Int'l Wrkshp on Research Issues in Data Engineering: Engineering e-Commerce/ e-Business Systems (RIDE'02)
1066-1395/02 $17.00 © 2002 **IEEE**

ner structure and operation of the involved providers. Hence it does not incur the risk of being unusable in practice due to technical mismatches and difficulties.

Finally, note also that 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.

5 Conclusions

In this paper we have considered certification issues arising in intragovernmental process supporting digital government services.

Certification of exchanged e-services is a critical activity in any application areas, but is of the utmost importance for e-government since very often exchanged data have a legal value and play a legal role.

The proposed solution, based on the introduction of two novel kinds of components, namely *dispatchers* and *network probes*, allows to efficiently and effectively deal with these issues. It also provide the technical platform on which to base control policies, if these are deemed necessary. Our solutions directly derives from successful experiences in the development of a number of real-life systems. Its conceptualization is described in [6], while further developments of the underlying formal model for data coherence maintenance among independent organizations are presented in [7] and [8].

Our approach requires a careful modification to be used in the case of public network, like Internet. In fact, very often in this case services are provided on top of a secure communication layer (e.g., SSL). The conceptual architecture here described can still be used, but radically different technical solution have to be implemented for the probes.

Acknowledgments. We would like to thank Elettra Cappadozzi and Aurora Girolamo for many useful and interesting discussions during the development of the work here described.

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., 46–56, Sep.99. Extended version published in the *Int. J. of Cooperative Information Systems*, Sep.01.

[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, 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, Preprint n.8/2001, Fakultät für Informatik, Universität Magdeburg.

[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] F.Arcieri, E.Cappadozzi, E.Nardelli, M.Talamo: SIM: a working example of an e-government service infrastructure for mountain communities, *Workshop on Electronic Government (DEXA-eGov'01)*, Conf. on Databases and Expert System Applications (DEXA'01), Sep.01, Munich, Germany, IEEE Computer Society Press, 2001.

[6] F.Arcieri, R.Giaccio, E.Nardelli, M.Talamo: A framework for inter-organizational public administration network services. *Int. Conf. on Advances in Infrastructure for Electronic Business, Science, and Education on the Internet (SSGRR'01)*, L'Aquila, Italy, Aug.01. IEEE Computer Society Press, 2001.

[7] F.Arcieri, E.Cappadozzi, G.Melideo, E.Nardelli, P.Naggar, M.Talamo: A formal model for data coherence maintenance. *Int. Workshop on Foundations of Models for Information Integration (FMII'01)*, 10th Workshop in the series Foundation of Models and Languages for Data and Objects (FMLDO), Viterbo, Italy, Sep.01. Lecture Notes in Computer Science Vol., Springer-Verlag, 2001.

[8] F.Arcieri, G.Melideo, E.Nardelli, M.Talamo: On the Dynamics of an Infrastructural Approach Supporting Coherence Maintenance for Inter-Organizational Collaboration, *Int. Symp. on Business Strategy Based Software Engineering (SoftwareTrends'01)*, Sept.01, Gersau, Switzerland, NetAcademy Press.

[9] F.Casati, U.Dayal, M.-C.Shan: Report on the VLDB Workshop on Technologies for E-Services (TES), *SIGMOD Record*, 29(4):11–15, 2000.

[10] A.K.Elmagarmid, W.J.McIver: The Ongoing March Toward Digital Government, Guest Editors' Introduction to the special section on Digital Government, *IEEE Computer*, 34(2):32–38, Feb.01.

[11] J.Joshi, A.Ghafoor, W.G.Aref, E.H.Spafford. Digital government security infrastructure design challenges. *Computer*, 34(2):66–72, Feb.01.

[12] Law Decree n.557 of 30/dec/93 and Law n.133 of 26/feb/94.

[13] M.Mecella, C.Batini. Enabling Italian E-government through a cooperative architecture. *IEEE Computer*, 34(2):40–45, Feb.01.

[14] E.Nardelli, M.Talamo, and P.Vocca. Efficient searching for multidimensional data made simple. 7th Annual European Symposium on Algorithms (ESA'99), Prague, Czech Republic, Jul.99, Lecture Notes in Computer Science vol.1643, pp. 339–353, Springer-Verlag.

[15] Simple Object Access Protocol (SOAP) 1.1, W3C, http://www.w3.org/TR/SOAP.

[16] M.Talamo, F.Arcieri, G.Conia, E.Nardelli: SICC: An Exchange System for Cadastral Information, *6th Int. Symp. on Large Spatial Databases (SSD'99)*, Hong Kong, China, Jul.99, Lecture Notes in Computer Science vol.1651, 360–364, Springer-Verlag.

[17] Bulletin of the Technical Committee on Data Engineering, *Special Issue on Infrastructures for Advanced E-Services*, G.Weikum (ed.), 24(1), March 2001.

[18] Web Service Description Language (WSDL) 1.1, W3C, http://www.w3.org/TR/wsdl.

Proceedings of the 12th Int'l Wrkshp on Research Issues in Data Engineering: Engineering e-Commerce/ e-Business Systems (RIDE'02)
1066-1395/02 $17.00 © 2002 **IEEE**