Handelshögskolan
VID GÖTEBORGS UNIVERSITET
Institutionen för informatik
2004-02-21

HOW TO SOLVE THE "MILLION USER PROBLEM"?

Identity and access management in a global environment

Abstract

Identity management and Access management are new concepts on the Internet, containing functionality as single-sign on and authentication. The need of those concepts is arising out of an enormous amount of possible users that are reached when companies put their core applications for easy access on the Internet. But clear signs of an early market make it hard for decision makers to find strategic solutions. Instead we have diverging visions from the market leaders and standard communities. New actors and solutions are still entering the market and challenge common used standards as LDAP and Active directory. One of these new challengers is the Baldo solution at Volvo AB. The Baldo solution is compared to two other ways of tackling the *million user problem* together with their advantages. For this investigation a number of interviews have been done with experts in the problem area. Can we create a solution that fulfil future needs of dynamic solutions and standards which can handle and build global solutions for a global market?

Keywords: Identity, Access, Management, LDAP, Internet, Portal

Author: Robert Brasegård (robert.brasegard@volvo.com)
Superior: Maria Bergenstjerna
Master thesis, 20 points, vt03

CONTENTS

1 INTRODUCTION

During my years working with computing and IT, *the million user problem* has been growing from a small problem to something that is in focus at the software industry. I work today as system responsible for a global Internet portal at Volvo Information Technology AB. The portal has today about 300 administrators in a global network. With this paper I will describe different solutions and its benefits considering my experience and knowledge about the problem area.

In the early 90's many companies developed an administration for users on a global market. But still all systems were locally installed on servers and PC's around the world. There was always a procedure to install, give instructions and to get everything to work. Installation was done locally and needed a physical presence of a person.

When new http-based clients on the Internet become a possibility, in the late 90's, this earlier hindrance was no problem. An explosion of new users and new applications has appeared. However the issue of handling authorization and roles for each user on each application has become a major problem.

The LDAP directory is a common used industry standard to handle users, groups and authorities (Howes, Smith, Good, Howes 1998). The use of standards is of course good and usable when they have been accepted of a large group of developers and companies. But in the light of the *million user problem*, is there a need to add functionality to this standard, to fill the gap of missing functionality on the Internet?

The things that happen on the Internet today tell us that traditional user administration is not enough any more. More aspects must be considered. Cultural and decision-making processes in different continents most be handled in the same system. As a consequence of this there is a need to build dynamic Internet portals that support different types of processes (Nakamura 2002).

The earlier *User administration* concept was not for a global environment as Internet; instead it was a facility to set authorization for users in a single application. A new expression, *Identity management,* has been founded. *Identity management* is a wider conception of what earlier was called user administration (Perry 2001). This is a composite expression and consists of two major parts (Yasin 2002). One is the authorization part, and the other part is the authentication which also contains the *Access management* concept. Uwe, Kreutzer, Zugenmaier [B] (1997) define Identity management as a concept that enables each user to express and to enforce his or her privacy and security needs in IT-systems depending on the situation the user is in.

What is included in the concept *Identity management*? On the Internet you can find different list of included concepts (Yasin 2002; Perry 2001; Jendricke, Kreutzer, Zugenmaier 1997). In this paper, the following sentences will be included in *Identity management* and *Access management.*

- Password reset
- Filter of information – Users will only see selected information.
- Authentication – password and user-id
- Authorization – what can a user do when he gets access to an application? (Access management)
- User profile – For example: Preferred language.
- Single sign-on

1.1 Background

Internet applications have much less possibilities to use sessions for user *access management.* This is a fact for all program developers that are moving from traditional windows developing to an Internet based application. Traditional windows techniques use sessions to handle authorities for users. The new web techniques can of course set a session id also but it will always lose contact with the server when the HTML-code is written at your browser.

An Internet expose will also reach an enormous scope of possible users to your portal. This will raise a number of new problems. How can for example 100 000 users be handled with correct authorities on 100 applications at your portal without an extremely expensive administration? The administration cost of an increasing number of users can not be underrated. This is a problem that all major companies are facing today when more and more business applications are exposed on the Internet. And a fact is that this enormous number of possibly users has never been faced before. Traditional companies like Volvo and many other big and middle size companies that will run a number of applications have different roles and authorities at those applications. I will give a comparison between three different solutions and how their way of tackling the *million user problem* will give effect on dependent systems.

A portal is a view which consists of underlying systems and information. The possibilities to create a system for a specific user are dynamic and unlimited concerning to information. The user himself will be something between layout designer and a user of information systems, connected to his id, done by different administrators. When portals are used for *single sign on,* users must have authorization at different levels for different applications. There is also a need to know how to administrate administrators. With this I mean that a super administrator can distribute authorities that are a subset of his own rights. This administrator can after that distributes subsets of his rights and so on. This will create a network of administrators, and it is a must, on a global market.

The portal will be dynamic and not designed by system developers. It is rather designed by the user himself and a number of administrators that probably not know the existence of each other. Variables which are not known when the portal is created will be easily added in a later phase. If the portal can be seen as a system of systems, many earlier difficulties in system design can be moderate. Instead the user will have problem with knowing what he/she shall add to the view of the portal and what he/she can demand from an administrator.

Web sites are now at the core of business interaction (Perry 2001). Companies use Internet for their core business and doing transactions with customers. This creates higher demands of design, system-architecture and user interaction.

A standard called LDAP, *Lightweight Directory Access Protocol,* was designed year 1992 at University of Michigan to adapt a complex enterprise directory system called X.500 and make it usable on the Internet. LDAP is frequently used on the Internet for handling users and their authorities and groups(Howes, Smith, Good, Howes 1998). But is LDAP enough? Is there missing functionality in the LDAP structure? The problem I will discuss in this paper is how LDAP and other additional solutions try to solve the above mentioned *million-user problem.*

We must build combinations of systems in a portal that match different cultures and processes globally. The core of culture is formed by values (Haynes, John D. Corbitt, Bryan. Theerasak, Thanasankit. 2001). Values are expressed as a symbolic form and include words, signs, colours and gestures. E.g. red colour can express feelings of very different kinds in the western society compared to the eastern countries. Symbols in China are very different in nature and meaning in comparison to their meaning in Europe. In China the dragon is a symbol used by the emperor and in Europe mostly a fairytale for the children. Of course this is not the same sort of dragon in China and Europe it differs in style and writings. The Chinese radiate power and authority and the European Childness and a bit fear sometimes. A global portal and its *user profile* handling must handle a variation of symbols, colours and other signs to handle a global market.

Figure 1 The Chinese dragon

A full-page advertisement in the *New York Times* (June 19, 2000) describes the impact of the on-line medias as follows.

Ours is the first culture in history to have moved inside media – to have largely replaced direct contact with people and nature for simulated versions of TV, sponsered by corporations. Now it's happenings globally, with grave effects on cultural diversity and democracy.

Today we can not se any human race on the Internet. Instead we can talk about *Cybertypes* (Nakamura 2002). Cybertypes are much more cultural rather than race related. Systems must be built considered to this. A global Internet portal must change considering who is connected to it and what intention the user has when he/she is connected to it.

How importance of trust must also be considered on the Internet are decribed by Uwe, Kreutzer, Zugenmaier (1997). Different cultures in separated parts of the world can have large different experience of trust by governance, state and other overall organisations. A user in a very controlled country must be sure that his/her opinions and actions on the Internet will not be logged and can be raised against him without a reason in the future (Fischer-Hübner, Quirchmayr, Yngström 1999). Other countries with large democracy and law systems may not consider this question as a problem.

There are two different types of identity on a portal on the Internet. One is added by administrators and system authorities connected to the user-identifier on the portal. The other one is social identities that are added by the user himself. An example is specific communities that the user has joined and added to the portal, if the portal gives this opportunity. A social identity can in this case be an Africa-Americans or Native-Swedish community or groups at Internet.

Demands on different organisations interactions with each other are in focus at large companies (Magoulas, Pessi 1998). To develop interactions between Customers, Suppliers and Subcontractors, for this purpose they have been started to use IT-technology. The solution for this is portals on the Internet. No other today known infrastructure can compete with this network. A fact is that you should have really good reasons for developing systems that <u>not</u> can be used on the Internet.

1.2 Purpose and Issue

Purpose:

This work will analyse IT-system needs of *identity* and *access management* on the Internet. I will analyse different ways to manage a growing number of possible combinations of user rights and application access possibilities when a portal is used.

Ways to handle questions as, how can an application be set to read only on the Internet, will be investigated. One example is, different ways to set authority at method and field level. Describe solutions that those can be handled and implemented in a cost effective way. Three different solutions are compared and benefits and shortcomes are described.

Target audience for this paper is experts within the *user administration* and *identity management* area. The study is a comparative analysis of three different solutions of an administration problem at a global market. I will compare a solution at Volvo AB with other commercial concepts at the market, at a solution based ground. Known benefits and disadvantages of each solution will be analysed. But still a much deeper analyse can be done. Therefore I will say that this investigation is only a light version and will only give an overall view of the problem. When I talk about a global dynamic environment I mean a global solution in a global environment, that is “think global”.

Issue:

- How can a growing number of users and applications be handled cost-effective and simple in a global dynamic environment?
    - What is missing in the LDAP standard to get a modern dynamic user administration?
    - Are the Identity management tools that exist today for handling users and applications enough granular for the user-administrators?

1.3 Delimitation

A major question of all companies on the Internet is of course how security will be handled, including firewalls, coding technique like *code behind* in the Microsoft .NET platform, but this paper will not handle that kind of questions.

Following issues will not be handled:

- I will not have any hypothesis about how system coding will be done to minimize the risk of being attacked from hackers on the Internet.
- Threat Management - How to handle questions about firewalls and what need of password security Internet is demanding
- What programming language that is to prefer.
- National-languages in portals and how to administrate translations on applications are not handled.
- Security groups that is normally and preferably handled by the firewall. Security groups are a facility that is often offered by more sophisticated firewalls and are a limitation of what applications you can reach. If you are connected to a group, then you can reach all application in it, if you know the URL.
- This work will not handle Identity management in other aspects than for access control and authentication in information technology.
- Identity management or Corporate Brand is sometimes used to describe how to create a distinctive image for a company. This work will not handle that type of questions.

1.4 Disposition

This paper has a comparative disposition who describes advantages and disadvantages between different solutions of the problem to administrate a major amount of users on the internet.

Chapter 1 Is an introduction that describes a background for this job. It also describes its purpose and issue together with its delimitations.

Chapter 2 Describes the scientific approach witch is qualitative. An approach like that will allow an iterative process of writhing. For this process a number of persons have been interviewed.

Chapter 3 Is the theoretical part. A number of theories and theorist have been used to make the *million user problem* understandable. Together with a description of common standards and their advantages.

Chapter 4 In this chapter tree different solution of the million user problem are described with their advantages and disadvantages.

Chapter 5 Describes those empirical tests that have been done. Positions at interviewed persons are described. What kind of portals and systems that are analysed. This chapter also analyses the *million user problem* in the light of interviewed persons. Phrases that have been told are rendered and commented.

Chapter 6 Is a description of the Volvo case that has developed the Baldo/Balda solution.

Chapter 7 Here is the summary of this work described. A discussion how different solutions and standards will effect the organisation if they are chosen.

Chapter 8 A short conclusion about the key issues in this work. What is the main task to get this problem solved?

Chapter 9 Have all references that are referred in the text. Many references are from the Internet and are in those cases connected to a URL. It also holds a list of figures and tables in this paper.

Chapter 10 Contains Codd's rules for how a relational database shall be build. Here are also the questions that are used in the interviews noted in chapter 5 and 6.

2 SCIENTIFIC APPROACH

A qualitative approach (Backman 1998) has been chosen for this paper. I have described how different solutions of *Identity management* will be apprehending of system developers and portal users. I have preferred a case study in the real life context and use of realistic phenomenon. This qualitative paradigm has a clear subjective accent but gives freedom to have an inference from other areas. This makes it possible to have an iterative writing process. A qualitative method describes solutions in a more verbal and experienced way. A question that is asked in the beginning can be modified and rewritten during the work. The case study process is from the beginning derived from medical and juridical cases, according to Backman (1998), where a case was synonymous with a problem or an individual.

A number of persons are interviewed and used as references to support conclusions. The key factor to select persons to be interviewed is their insight into the problem area. This is not always the fact, but the intention has always been to catch that key person. The selection of interviewed persons has been done considered to the company's knowledge of *Identity management* and their position at the company. The questions that have been used in the interviews are included in appendix 8.2. In the interviews, a semi-structured interview technique is used, which make a more open dialogue between the interviewer and the interviewed. The focus of this qualitative interview technique is not to get structured results as tables (Jacobsen 1993). Instead new things that occur in the interview can be followed up and illuminated. This technique uses a type of questions that are called *open questions*. The interviews was partly recorded and later written to paper and sent to the interviewed for comments. The main focus at the interviews was to catch new thoughts and ideas of the identity management area. One other ting that was the purpose of the interviews was to locate if there is anything lacking in the standards that are used by the industry today. And if it is, then what can be done to fill this gap.

This field of questions concerning *Identity management* and *Access management* is a new problem area and very few literature references can be found. However a number of white papers are written and can be found on the Internet. These papers will be studied and referred to. I will also make references to traditional literature.

2.1 Interviews

To create a knowledge ground for my investigation a number of persons have been interviewed. I have interviewed 5 persons outside the Volvo organisation to create a knowledge base of the problem area. Later was one manager at Volvo Truck International division interviewed for specific Baldo questions.

Descriptions of persons who have been interviewed are listed below:

No	Position	Type of company	Experience
1	Market responsible	Participant in the Identity management sector	Long experience in the problem area. Knowledge about PKI infrastructure and the LDAP and SAML standard
2	Executive Director	Participant at the automotive industry	Limited experience in the problem area. Knowledge about EDI, IDN and X.25 standard.
3	Executive Director	Participant in the Identity management sector	Long experience in the problem area. Knowledge about the LDAP standard.
4	Executive Director	Participant in the Identity management sector	Long experience in the problem area. Knowledge about Access management and Identity management.
5	Consultant within the Identity management area	Technical Consultant	Experience in the problem area. Knowledge the LDAP standard, directory servers and Meta directories.
6	Manager at Volvo Truck International division	Volvo Trucks	Experience in user administration area

Table 1 Interviewed persons

The interviews outside Volvo have been done by telephone. The reason for this is geographical. It has not been possible to travel the distance and do the interviews on place.

The purpose with the interviews has been to catch trends, thoughts and visions of the interviewee. The questions have been open and semi-structured; whish creates a way for the interviewee to answer as free as possible. The questions have been more as main points for a communication between the interviewer and the interviewee. This technique demands a high level of insight into the problem area by the interviewer. One disadvantage with this method is that is difficult to put the result side-to-side and analyse it. But a qualitative approach of the interviews has, by my opinion, been the most appropriate for the type of investigation that is done in this paper.

I have used 10 questions (Appendix 9.2) where 3 concerns Organisations, 3 Design and standards, 3 Culture and Users and 1 more vision and future oriented. Those questions are done for interviews outside the Volvo organisation.

For the interview inside the Volvo organisation I have used the 7 questions in Appendix 9.3

3 THEORETICAL VIEW

My approach have been to lift the *million user problem* from a very technical level to a more administrative and conceptual level. For this I have used an Infological approach described by (Langefors 1995). Langefors summarizes the infological approach as follows:

The Infological approach was based on the observation that the users should have real control of the system design and that this could be made possible by exploiting the fact that the main system design is an organizational design and that the needs analysis can bee free from technological aspects and language.

Langefors also says: Good technology requires not only efficient design and technical quality. The technologist must be interested in contributing to the well-being and happiness of the clients. This has two aspects:

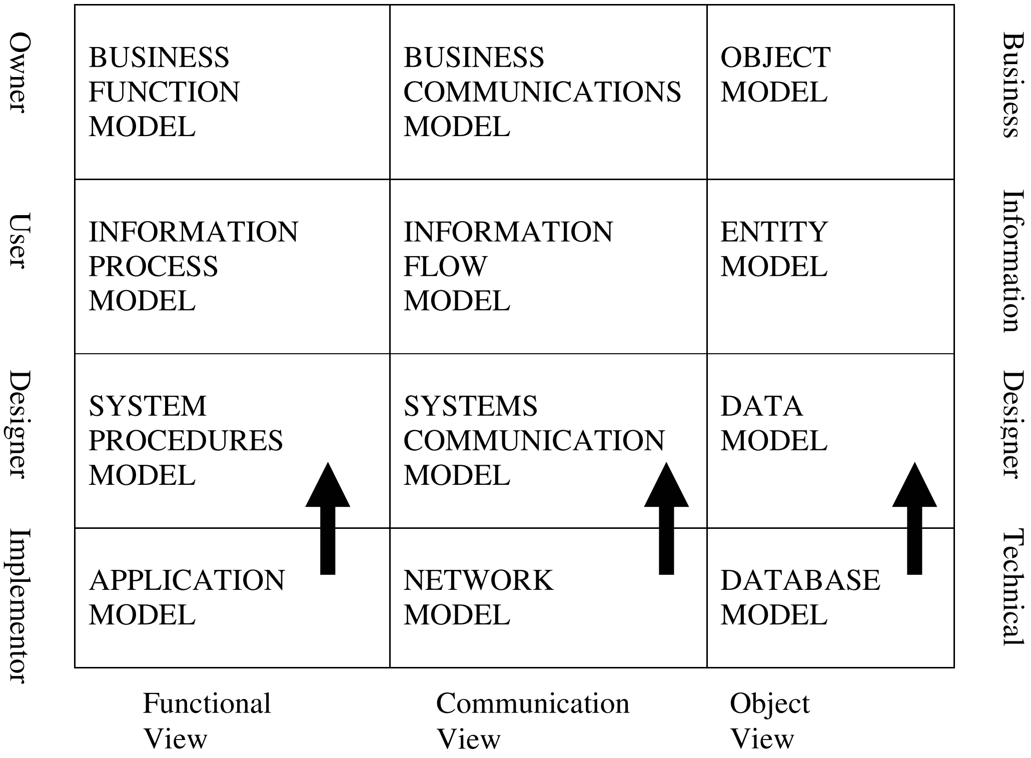
1. To satisfy needs or desires the client is aware of and requests to have satisfied (the market).
2. To create new desires, unknown to the client but made desirable to him, and develop solutions to these (create a market)

The system design approach described by Langefors about *complex systems design* (Langefors 1998 page 67) is as follows.

Complexity is the property of being a thing that can only be perceived piecewise. Thus, to understand a complex thing as a whole, we have to study it as a system; that is, we have to examine the system formed from the perceptible pieces through analysis of the interrelations among those pieces. The effect of an event in one piece spreads to other pieces through the relationship among them.

Langefors also talks about *workability*. This is a way of breaking down a global goal to more local and understandable goals, so the user and administrator can understand it. If the global goal is only known by the user himself, local goals must be understandable.

The analysis in this paper will move the perspective from most machine close level Technical/Implementor upwards against the Designer level (van der Poel 1989). The traditional handler of questions like *access management* is the technical/security manager. I will move this to a more administrative level (Figure 2).



Owner	BUSINESS FUNCTION MODEL	BUSINESS COMMUNICATIONS MODEL	OBJECT MODEL	Business
User	INFORMATION PROCESS MODEL	INFORMATION FLOW MODEL	ENTITY MODEL	Information
Designer	SYSTEM PROCEDURES MODEL	SYSTEMS COMMUNICATION MODEL	DATA MODEL	Designer
Implementor	APPLICATION MODEL	NETWORK MODEL	DATABASE MODEL	Technical
	Functional View	Communication View	Object View	

Figure 2 Perspectives and viewpoints for information planning (van der Poel 1989)

If the administration of users, authorities and their access-management is moved from the lower level to a higher level in the van der Poel model above, we will create an easier and more understandable solution for the administrators. The level whish will solve the *million user problem* is probably the Designer level, but a move to an even higher level is preferred.

The computer network theory as described in the book Computer Networks (Tandenbaum 1988) will be used for describing the design of the Baldo/Balda concept. When I talk about computer networked system I do not mean centralistic systems or MIS-systems (Davis 1985). Literatures distinguish between computer networked and centralistic (distributed systems), is that is computer networked systems have its own conceptual and database model, with public API:s for interactions with other systems. Centralistic design creates a global object model which is including all objects and attributes of interest. My point of view is that centralistic systems also called Management Information System (MIS) will not be built today and will not be discussed more over in this paper.

For database normalizations and general design analysis, I use traditional relation database theory (Date 1990) and I will in an extended thinking of this theory use a notation that can be called typed objects. When I talk about typed objects, I mean design of the database that will minimize database changes when new attributes are added. This is done by letting a table grow vertical instead of horizontal when a new attribute is added. In system maintenance of information systems, a database built with typed objects will facilitate changes. A fact is that database changes are the most difficult changes to handle in system maintenance and must be avoided if it is possible. A database change will also always demand a new release of the system. If typed objects are used, most attributes can be added without a database change and often also without a new release of the system.

A useful way of defining the logical form of a database is to use a relational model of data as described by Codd’s rules (Appendix 9.1) this is also described by C. J. Date 1990. Using the relational model instead of a hierarchical database allows the relationships between data items to be defined without considering the physical database organization (Sommerville 1989). Arguing about this should not be necessary in year 2004.

At analyse of *identity management* (Yasin 2002) there are distinguish between Authentication, Access control and Authorization (Baltimore 2001). Authentication is a process that identifies a user individual through a username and password. This process is often and preferably done by the firewall and therefore outside the limitation of this work. Access control gives a logged-on user access to specific applications on a portal, e.g. can a user get five of ten possible applications at his view of the portal. An application is defined as a provider of various means of recording, searching, extracting and distributing information (Gunton 1989). A portal consists of a selective number of possible applications and more static information-pages. Authorization gives the user different levels on possibilities to make actions of his accessed applications, e.g. *read-only* at application 1, *administrator* at application 2 and *super-user* at 3 and so on.

3.1 Existing standards

Some of the standards that are used today are transformed from early EDI-standards as the X.400/X.500 standard. Other standards that are entering the market are immature, and show signs of an early market (Dalton 2001, Wainewright 2002).

An OASIS standard, Web Services for Remote Portlets (WSRP), aims to make it easier to automatically include services in portals. If this standard will be common used it's a promising tool for plug-in functionality.

3.1.1 LDAP

Is the LDAP- standard something that can take care of this global demands or must there bee additional services? The earlier X.500 standard was too big to be adopted of the Internet. The LDAP directory is a hierarchical database that has the same disadvantages and advantages that hierarchical databases always have had. A hierarchical database is always rapidly fast but has failings to handle multi dimension questions. In figure 6, the LDAP directory is described as an *early binding policy store*. LDAP is today the most common standard on the Internet for handling authorities.

If LDAP was used for the *early binding* (see 3.4) and complemented by a relational database (appendix 8.1 Codd's rules) for the *late binding* many demands and difficulties disappear. LDAP is often described as a telephone book, for example the white pages. But the problem is that on the Internet will handle questions like selected users in white pages, compared with yellow pages including filters and belongings at the same time. But LDAP is a fully accepted standard and will work very well for the *early binding*.

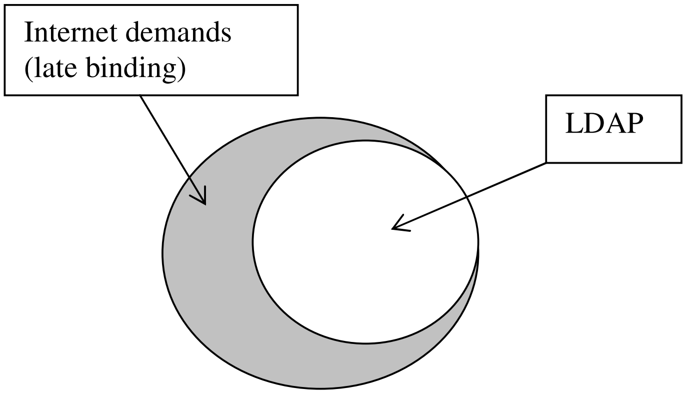


Figure 3 LDAP vs. demands on Internet

One advantage with LDAP is of course that it is a standard. This is, and will always be, a big advantage. But at some point a standard must be replaced by a more up-to-date standard.

LDAP have **<u>not</u>**:

- transactions
- optimizing for updates/deletes/adds
- model (relational model)
- stored procedures

3.1.2 IETF (The Internet Engineering Task Force)

The Internet Engineering Task Force (IETF) is a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet.

Here is a discussion about why or why not a company should use LDAP in their email solutions. Email solutions are one of the major areas where LDAP have a traditional large number of instances.

Why not use LDAP?

- *Protocols are most successful when focused on initial design goals. LDAP was designed for global, authoratative user directory, not ACAP/IMSP design goals.*
- *Protocols get more complicated and less implementable as features are tacked on. We don't want to contribute to creeping featurism in LDAP as it is already useful in email world to solve the global authoratative user directory problem.*
- *LDAP doesn't have experimental background for addressing the requirements of ACAP/IMSP. The LDAP model is sufficiently different from IMSP that it could create unforseen problems.*
- *Deployed LDAPv2 infrastructure is missing fundamental ACAP requirements. Upgrading a deployed and much needed infrastructure is far more expensive then adding a new protocol.*
- *LDAP's binary X.500 based model makes the protocol difficult to debug. The general application configuration problem needs a simpler protocol if it's going to end up in many applications.*

(The Internet Engineering Task Force [online])

Other discussions about how to handle questions about problems that is not handled by LDAP is described widely on the Internet. One other approach is the IMSP/ACAP standard that is described below of the University of Maryland. IMSP/ACAP is a standard to handle email solutions.

IMSP can be looked at as a "preferences server". With the realization that users don't always read email from the same machine, IMAP was created to allow users to manage all email on a server. However, it quickly became apparent that users are still somewhat tied to a particular machine, because the preferences for their mail reader were stored on the local hard drive, and because of that, they had to re-setup their mail program for each new machine they use, which was a real pain. Because of this, IMSP was created - basically, with IMSP, you can tell a mail client who you are and where your IMSP server is, and it retrieves all your preferences (windows locations, subscribed folders, etc) from there. Once IMSP was used for email, it was realized that with a little generalization, this would be useful to many programs other than just email, so ACAP is being developed to replace IMSP as a general use preferences server.

(University of Maryland Biotechnical Institute [online])

LDAP and IMAP have their origin in this e-mail type of solutions. A transform to the Internet *million user problem* is maybe not so easy.

3.2 Signs of an early market

Web sites are now at the core of business interaction (Perry, Robert. 2001). As the web has matured, the expectations of web providers have increased dramatically. Users demand easy access to core applications at companies for fast access to critical information. Content at portals must be updated rapidly and documents must be visualized and removed on demands. Personalization has become one of the most important aspects on portals. With personalization the portal change its way of viewing information for each user that is accessing the portal. Other aspects are *Multiple Web destinations* and are a way to handle multiple media for presentation. E.g. Low bandwidth, broadband, PDA, phone and so on. This will create high dynamic solutions on the portal that will adjust for receiving media.

Sites are big and getting bigger (Dalton 2001). But today's products are immature, and show three classic signs of an early market. To illuminate how an early market makes it difficult for decisions makers to get good information for their decisions.

1. ***Incomplete products***. Though most products demonstrate proficiency in one or more categories. E.g. security or performance.
2. ***Poorly defined category***. There exists a basic problem that companies do not agree about what different statement means. A definition of statements must be established.
3. ***Diverging visions***. Tool providers generate equally conflicting roadmaps for the future.

Content managers of a portal need a long term plan, but an early market can not give them this. They have also limited experience of portals and to administrate this number of users. Therefore most companies choose a solution but can not really decide if it is a good investment or not (Dalton 2001). A wrong investment can give the following problems:

- ***Users can not find information.*** Pore possibilities to filter information to end-users make it difficult to find the right information. This type of problem everyone can find, when searching on the Internet.
- ***Process anarchy.*** On a global market many processes must be handled. A portal that not has dynamic possibilities to adjust itself for different processes in different areas of the world will run into major problems.
- ***Integration woes intensify.*** As the number of content formats and business initiatives accelerate, raises a need of standards for new products. But they will find that this kind of standards is hard to find.

Today (2003-03-01) more than 80% of all companies have a portal. Most common is portals for employees 64% followed by customer portals 49%, distribution 29% and supplier 25%. Today Oracle with 49% of the portal-market is the biggest participant. One reason of Oracles big share of the market is that they offer their portal tool as a free addition to Oracle program server. This type of accessing the market like Oracle use is called *bounding*. Other big participants are Peoplesoft, SAP and IBM. The market today show clear unripeness and consolidation is expected, according to Wallström (Computer Sweden no 19, 2003).

3.3 Approach to structure parts in the problem area

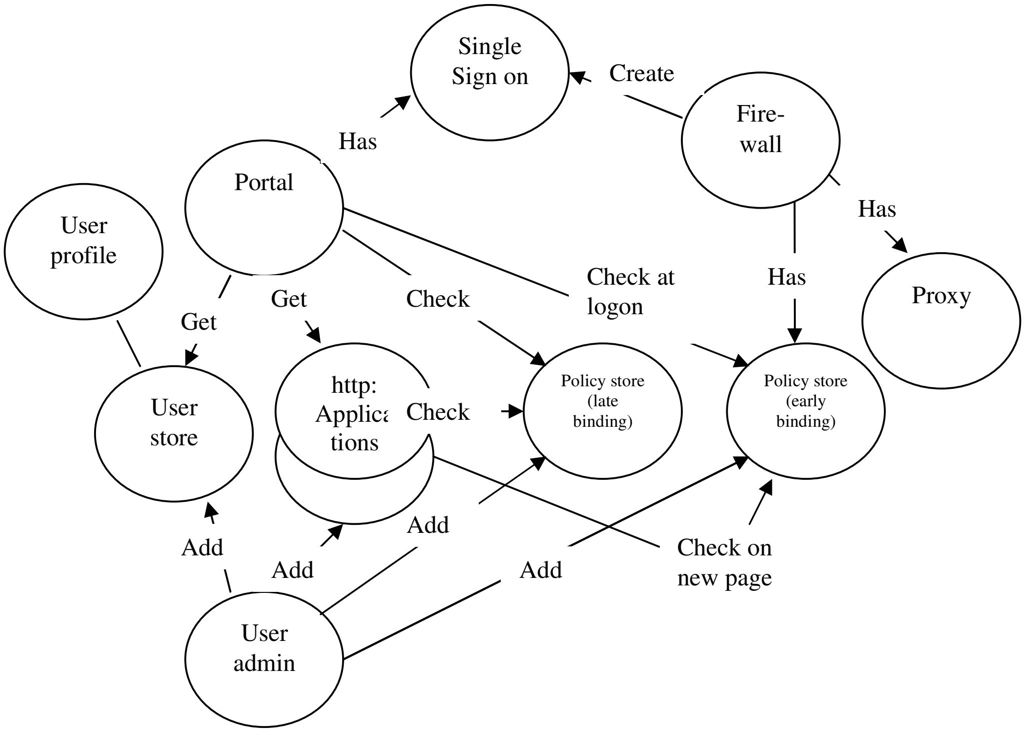


Figure 4 Interactions between parts of the problem area

Figure 4 describes whish parts that are connected to *the million user problem* and how those parts interact with each other. The firewall consist of at least an *Early binding policy store* and a *Proxy* (most firewalls offer much more facilities). Where the Proxy optimise the transactions and the *Early binding policy store* checks if a specific user have access to a specific domain of web-pages and in some cases also set roles and attributes for the user. The *Late binding policy store*, in this case, gives instruction to each http-page that subscribes on this information and can be e.g. enable fields and buttons, give information filter for this specific user and other user specific information. A User profile exists and controls attributes as preferred language and background picture at the portal.

A very important aspect is to reach Single-sign-on for a user that has logged on at the Portal via the firewall. The single-sign-on are created when the firewall give the user access via the *Early binding policy store* where his/her links are stored and referred to. The single-sign-on concept have for a long time been at vision for traditional developers, but have first trough internet been a real possibility.

3.4 Early Binding vs. Late Binding

Early binding is a solution where all permissions, roles and authorises are defined in advance (Gebel 2003). This is often done already when a user login to a portal or system. The *policy store*, in an *early binding* solution, set a cookie that includes all roles and permissions for the logged on user (see figure 6 The firewall solution). This cookie is thereafter sent to all links that are connected in the portal or system and can be read by anyone that are interested of the information. E.g. what links that is allowed to be accessed for a specific user. The PKI *Public-key Infrastructure* technology is an example of *early binding*.

Late binding permits more dynamic actions to be taken during processing instead of relaying on the ability to map out every possibility in advance says Gebel (2003). Rule-processing engines examine and evaluate user attributes and makes decisions “on the fly” (see 4.3 The Baldo/Balda solution that combines early and late binding).

3.5 Portals

How much functionality must be included for a link-page before it can be called a portal? If a site only contains links to other applications it can not be classified as a portal.

What is a portal then? The word was first used to describe the sites of popular Internet access providers or search engines such as AOL, MSN and Yahoo! A portal is a web access point. And it’s consists of web pages that act as a starting point for using the Web or web-based services. But this is not enough to be called a portal. It must be a corporate portal or enterprise information portal that acts as a starting point for employees or associates of an organization to access corporate information and multiple applications in different environments. Portal software has become a distinct class of web server software that acts as a platform for deploying portals (Wainewright, 2002). Connected to this we must have a *late binding* rule access that gives a logged-on user access only to selected services and selected information on each of those applications.

4 THREE DIFFERENT SOLUTIONS

In the light of the sentence *whole is more than the sum of its parts* (Platon and Aristoteles) I would say that a portal and *identity management* has its focus on the *whole*. At a portal the whole is built of different administrators, maybe globally, and they will not know the existence of each other. Together these administrators are building a *whole* for one user without knowing and understanding its parts. The *whole is more than the sum of its parts* will only be adopted by the user himself. No other administrator will understand and even know what the *whole* is for a specific user. The user can often also connect his personal links, as personal email links, to the portal and make a personal look at the graphic layout. About this no one knows, except for the user himself, but the sum will be far more valuable for the user than the parts it consists of.

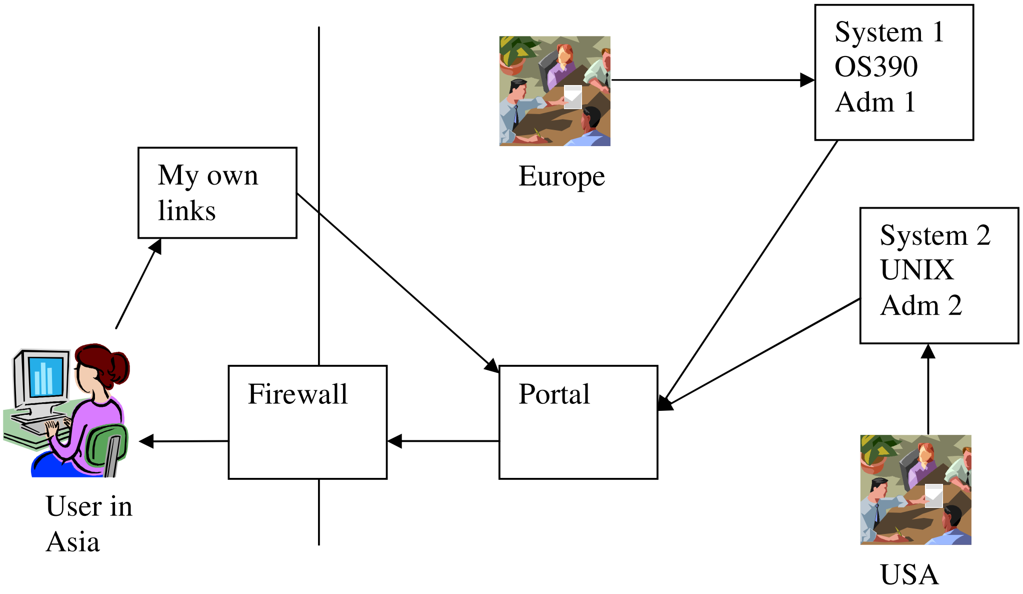


Figure 5 Whole is more than the sum of its parts

4.1 The firewall solution

Traditional Security Companies work with a module connected to the firewall called *policy store* (early binding). The *Policy store* contains domains, roles and sometimes categories and actions. This is nearly always handled by the LDAP standard, but other solutions can be used.

Domains are a part of *early binding* and are a hard firewall limitation to access an URL outside the domain. Domains and role based security is sometimes called *Security Control Centre*. This will prevent users to access an application outside the domain even if he knows the complete URL and this is a necessary security facility on a firewall. Some more integrated tools also offer more application usable possibilities like roles, categories and actions.

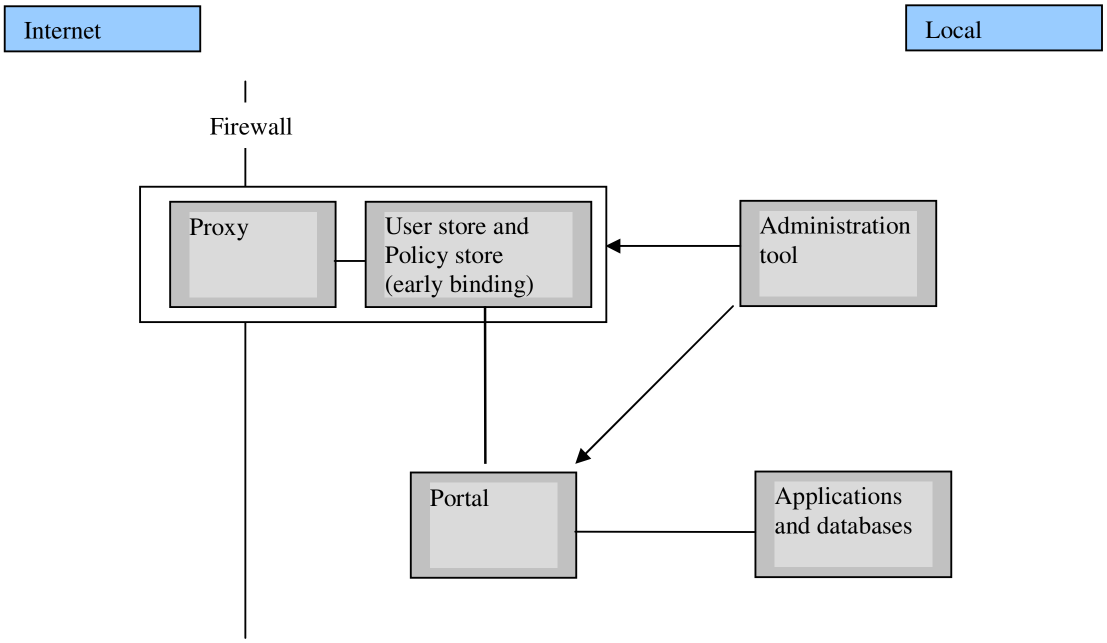

Internet
Local
Firewall
Proxy
User store and Policy store (early binding)
Administration tool
Portal
Applications and databases

Figure 6 The firewall solution

The Proxy, User store and *Policy store* is connected to each other and will handle all user attributes. When you connect these two modules you will put all user attributes at a technical level.

The LDAP standards have no natural administration facilities. For this an administration tool must be used. The administration tool has a central function in all examples in this paper. An effective administration tool will reduce cost and give administrators a better work situation. The portal is in figure 6, is any portal that is used by a company.

This solution often use a technique called PKI and this is a way of put all user information in a certificate. All applications that want to find roles and other settings for a user, can find the information in this certificate (Baltimore 2002). It works as a cookie that will be set and used for all http-pages that have a need to use the information.

To choose a solution offered by traditional security companies, in this paper called *the firewall solution* (Figure 6), is today one of the most common, *Identity management* market leaders are often using this solution. This solution will offer a possibility to add all LDAP information of a user together with other Login information, as *logged on user*, into the http header. This technique is sometimes called *PKI – Public Key Infrastructure*. This is a technically complex solution that offers wide spectra of different opportunities of using a structure known by the company.

One disadvantage of this solution is the setting of categories and actions at applications in a portal. That is a result and effect of the hierarchical-structure in the LDAP standards that is used for handling the *policy store*. A role, category and action handling is very complex and will be difficult to handle in a tree-structure, even if searching in trees is a very fast accessing method. When this method to give attributes to a user is used, large quantities of information must be shoved between http-pages. It can be a problem for some web-servers to handle this large amount of data, especially at clustered solutions.

The benefit is that all user information can be found in the header without accessing any database or directory. All pages that need to know roles, authorities or user attributes have them and can use them directly.

Most security companies work with an integrated authentication and authorization and it works in many cases like in the figure 6.

4.2 The database solution

Database solutions of the *million user problem* often use triggers on the database level to set filter and other user-specific attributes. One advantage this solution offer is a high security. By using the database as *policy store* all database securities and facilities can be used to secure actions in the system.

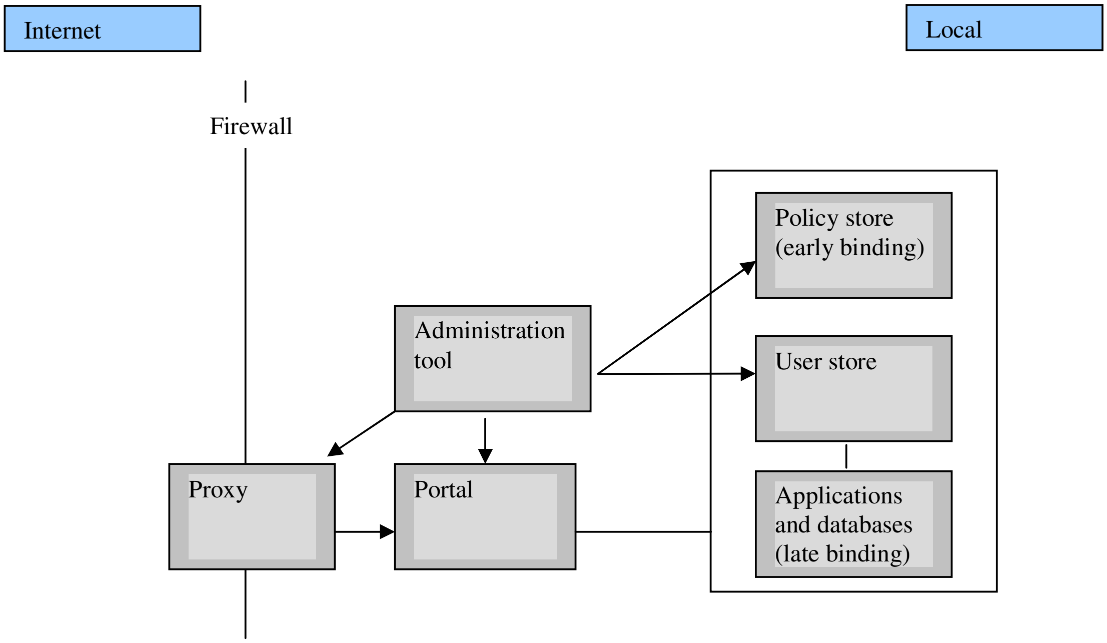

Internet
Local
Firewall
Policy store (early binding)
Administration tool
User store
Proxy
Portal
Applications and databases (late binding)

Figure 7 The database solution

A disadvantage is that all users have to be added as users in the database, even if they never or seldom access the database. This will create big and expensive license costs (The Fact Point group 2003).

The database solution does not need, but can have, LDAP standards in the *early binding policy store*. And the LDAP standard is also for database solution the most common solution. But to use LDAP for the *late binding policy store* will not give the flexibility that are demand for that functionality, instead this solution use the original database security for the late binding.

Database/software companies usually put the *policy store* connected to the database. This is of course more in line with their interests (Figure 7). Using database roles and user-access offered by normal database-administration will solve most problems at Internet *identity management*, but there are several aspects that talk against this solution. One is that it will be very dependent on one specific database. Another is that you have to register all portal-users as users in the database, even if they seldom or never access the portal. *Policy stores* connected to the database will also be hard or impossible to adjust for applications in other environments.

Oracle, who is market leader in the portal area, have a big advantage (Wallström 2003), when new groups of users who already use Oracle products are choosing a Portal development tool. It is an easy way to get a fast result for developers that already are used with Oracles development tools.

4.3 The Baldo/Balda solution

Categories and actions are normally handed over to the applications themselves. It can be hard to administrate, e.g. 100 applications in different environments AS400, OS390, UNIX, Windows and so on. The traditional way to administrate applications, by letting them administrate their authorities themselves, is not administrative acceptable on the Internet. The traditional solution will require a lot of man-time and economic resources because it is complex to administrate. This is the background for the Baldo/Balda solution.

The solution use LDAP for the *early binding user store* at the firewall. This with a combination of *early* and *late binding* are used where a relational object model is used to hold all user profiles, authorities and filters in the *late binding*. The benefit of this solution is an easy and dynamic administration of all necessary attributes for a user in the *late binding* and a high security offered by the firewall trough *early binding*. The relational model will also contain fine granular information for each *HTTP link* in a portal and this will make it possible to set authority on method and field level for each user.

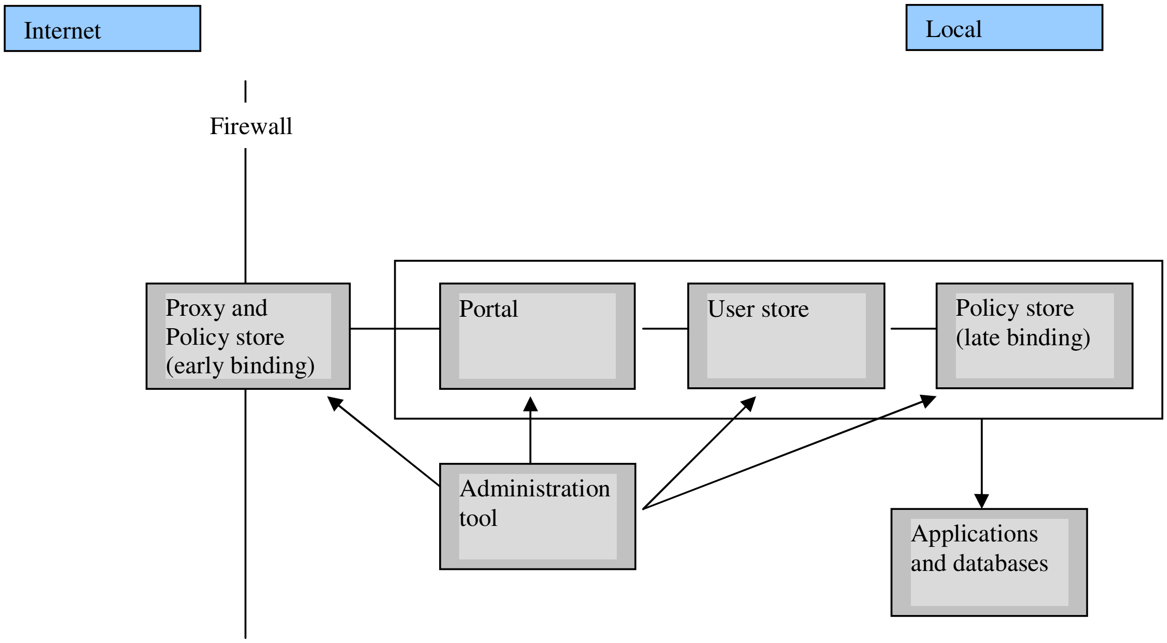

Internet
Local
Firewall
Proxy and Policy store (early binding)
Portal
User store
Policy store (late binding)
Administration tool
Applications and databases

Figure 8 The Baldo/Balda Solution

Roles are related to the user and will place him in his special area, e.g. a Car Dealer. Categories will be connected to a specific application and give a user a user-category on the application, e.g. read-only. Actions will make it possible for a user to perform action on each category on an application, e.g. delete, save or update.

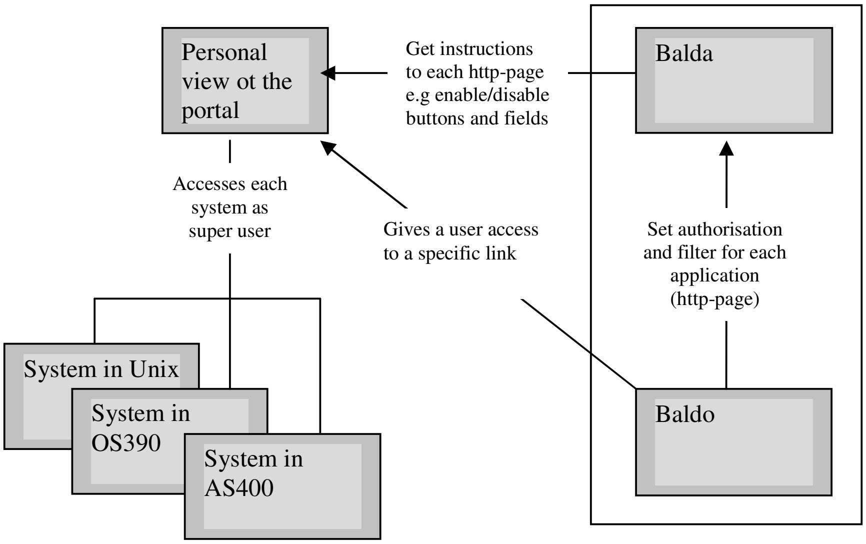

Personal view ot the portal
Get instructions to each http-page e.g enable/disable buttons and fields
Balda
Accesses each system as super user
Gives a user access to a specific link
Set authorisation and filter for each application (http-page)
System in Unix
System in OS390
System in AS400
Baldo

Figure 9 Baldo identity management in a multiple operative system environment

A positive consequence of the baldo/balda solution in figure 9 is that you can run all your users on a single user, the super user. The http-client get all instruction how to enable or disable buttons and fields, set filter on information, whish roles a user has and so on. This will of course demand that system responsible follows design guidelines for this solution. A strategic decision must be taken from the IT-governance at the company. This will of course only have affect on “home built” solutions, where this can be controlled. On bought applications, e.g. SAP/R3, authorisation can be set local and only the link to the portal will be controlled. This will create a combination of “home built” and bought application that can be accessed in the same portal. This solution will also create single-sign-on for all application that are built with the concept.

Dynamic links in the portal use the *late binding policy store* to get all actions on each link. The first action every link must do, is ask the *late binding policy store* about what *actions* logged-on user has on activated http-link. E.g. the portal asks the *late binding policy store* with active user-id as parameter. Answer can be: enable add-button, enable delete-button, and enable field1 and so on. This will get the advantage of that the database can be accessed with the same user every time, because the http-page always gets instructions from the *late binding policy store*. Only the sever machine will set the limit on how many sessions it can have open at the same time. This is also of course a large benefit considering licenses when only one user/password is needed for accessing the database.

A benefit of this solution is that all functionality offered by *the firewall solution* is not needed. What we need in this solution is the main password/user-match together with security group offered by the *early binding* at the firewall. This will give us possibilities to select solutions that are less expensive.

The “hard security” checks are done by the Firewall and the *early binding policy store.* This put high demands of the firewall security. But a fact is that this is exactly why the firewall is good at.

A user administration tool for updates and deletes in both applications connected to the portal and both *policy stores* is necessary for an easy administration. It is also a benefit if this tool can be used to add links to a specific user-id on the portal.

5 EMPIRICAL VIEW

For empirical testing an Internet portal called Technical Support Centre (TSC – main site) at Volvo Parts in Gothenburg will be used as a reference. This portal has underlying systems for handling user administration and security. At this portal Baldo is a user administration tool and Balda is a security engine. The Baldo solution is compared with two other common solutions on the commercial market.

BALDO/BALDA is today in use at Volvo AB and handles about 36.000 users in 130 countries and is used as an empirical test for the Baldo/Balda solution. Baldo is a user-database originally built in traditional windows technique, built for Volvo Parts with a start around the year 1997. Baldo was rewritten around the year 2001 to an HTML-based Internet application and is today generally accepted at Volvo Truck as the User administration tool for dealers and partners. This growth and acceptation can be explained by the security engine called Balda (late binding), being developed when Baldo was rewritten. I will describe benefits of this solution of administration and security compared with other solutions of security engines.

The Software industry has today focus at *Identity management*. But a fact is that traditional companies which have handled *User administration* have a big advantage against new participants on the market. Companies with this background have for years handled business needs as authority, filters of information and actions on applications.

LDAP is a directory that runs at servers at host computers on the Internet, and various client programs which understand the protocol can log into the server and look up entries. The complete X.500 is too complex for a regular PC and to be handled on Internet. LDAP servers exist at three levels: There are big public servers such as BigFoot and Infospace, large organizational servers at universities and corporations, and smaller LDAP servers for workgroups. LDAP servers index all the data in their entries and *filters* may be used to select just the person or group you want, and return just the information you want. This LDAP solution uses tree-structure to organize the structure for the user handling. For an example, users can be placed under a country, city and district in the directory.

5.1 The million user problem

The *million user problem* must be moved from a technical level to a more conceptual level. This process of trying to move technical complex solutions to a higher level in van der Poel's model (Figure 2) is an ongoing process, and it will always be an ongoing process, when new technique are getting more and more used by a lager group of administrators. The problem is not only to find a solution, it is also, and maybe this is an even bigger problem, getting it accepted and making some kind of commercial product or standard of it. Below phrases and statements from the interviews have been put into the problem areas in this paper.

- *Answers from the interviews are in italic letters.*

If the *million user problem* will be moved from the lower levels of problem solving areas in Van Der Poel's model (Figure 2) to a higher level, will this have any effect on the way of tackling the problem? Of course this will be a translation problem, as it always is, if something gets more conceptual and understandable by larger groups of users. There will also be a new way to look at the problem.

- *There is today a trend…that big user groups is connected to more and more web sites*

A more conceptual and administrative way of tackling the problem is necessary, if a user administrative hausse is to be avoided. The amount of users that are connected to the Internet has never been seen before.

5.2 Organisation and market

- *The market shows signs to be an early market… what is needed are standards that will make it possible to buy an application… with plug-in functionality*

To reach a point of plug-in functionality there must be an industry standard that is generally accepted. The fact that the market leaders are using different solutions and diverging visions is a sign that we have not reached that point yet (Dalton 2003, Wainewright 2002).

- *There is today too many participants on the market…joint ventures are to be expected*

Clear signs of an early market make it difficult for decision makers to get good information.

- *It is theoretically difficult to create business case…and it is hard to find success stories…*

Multiple environments make it difficult to create business case. This was an opinion from one interviewee.

- *This early phase of development has no really good success stories…*

This shows that an early market and development of standards and solutions will continue in this area.

- *We can today talk about id-creators and id-consumers at the internet…*

That is an interesting view point, to see other system that are interested in id-information as roles and actions, could be seen as id-consumers.

5.3 Cybertypes and cultures

There is also a different way to look at IT in different countries. The Company culture and the way to look at the decision-making process have a big influence. Therefore a portal must contain possibilities to change its behaviour.

One interviewed participant in the automotive industry says:

- *Some countries only look at IT as an advanced typing-machine… and others… use IT to reach a high level of automatization …*

- *Differences between users…need a flexible way of handling different cyber cultures…*

The *identity management* concept must in its user profile parts contain a possibility and flexibility that allow users to have different views and adjustments to its *Cypertype*.

- *When I worked in Belgium…the employees had to apply for access to the Internet from their employer …*

This shows how different employers in different countries can look at how to use the Internet.

5.4 LDAP and other standards

All participants in this problem area have been in contact with LDAP, which are the most common solution and standard. The interviews have nevertheless shown missing functionality in this standard. The LDAP standard is a hierarchical structure and has lacking functionality in data relations and it is often used to solve problem that it is not made for.

- *LDAP is only a catalogue protocol to handle user attributes...*

It is important that LDAP is not used to solve problems that it was not designed for. Companies that try to solve all problems with the LDAP-standard have not the flexibility that is needed for the *Identity management* concept.

- *LDAP is extremely fast to fetch information about a single user…but is no good at updates or deletes.*

- *There is today lacking functionality…in common standards…*

- *A new standard…SAML…is in use*

SAML which is an XML framework for exchanging authentication, attribute and authorization information, is designed to enable secure *single sign on* to applications within organizations, as well as across companies. SAML is a standard for authenticating and exchanging user identities across Web applications and services. But *single sign on* is only a small part of the *Identity management* concept. This standard will not help us to handle *the million user problem.*

- *We are using the new standard for Windows 2000… Active directory… to handle those questions…*

A standard that is used by Microsoft's Windows 2000 is called Active directory. This standard is often used to handle printers, workstations and other recourses in a network. It has also been used for *Identity Management*, but this will not be a good idea. Directories like LDAP and Active directory will not be able to handle the questions that are declared in this paper.

6 THE VOLVO CASE

Year 1998 Volvo started to develop a firewall solution by them selves. This was a solution very similar with technique offered by commercial solutions today (2004). But a decision at Volvo IT governance decided that this kind of technique should be bought and not developed internally. After this decision a new firewall was bought and the choice fell on the commercial product Siteminder, who is a market-leader in this area. And today Volvo use two firewalls, one internally developed and one bought on the commercial market. A single-sign-on between these products will not be developed by Volvo.

Around year 2001, when Baldo was designed, very few white papers and other articles were to find, concerning access and identity management. The way of combine *early* and *late binding*, was therefore developed only from needs at Volvo. At this point, no other known solution was been designed like this. At least there was no information or white papers to find that described it.

During year 1998 at Volvo, there was also a need of a user administration tool mostly for their Truck programming tool VCADS Pro, whish is used for programming software in Trucks and other Volvo products. The separately developed user-administration-tool, called Usams, controlled users and their authorities for programming trucks and other products with VCADS PRO.

Year 2001 a portal for Volvo Parts, called TSC-mainsite (**T**echnical **S**upport **C**entre), was created. But no good access and identity management tool was to find. At this point Usams come up as an alternative to control even authorities at internet. A work started to investigate if this was possible. Quite soon it was obvious that major changes of Usams had to be done, and a there was a need of a new concept to control many tools at the same time, not only VCADS PRO. Still the information and experience of Usams was a ground to build on.
A decision was made at Volvo Parts; this was the way to go.

A new name was created – Baldo (earlier Usams), and for the security module – Balda (the *late binding*). The focus was to have a dynamic solution whish could interact with any firewall and it's *early binding*. And the *late biding* should be offered to any software-tool who wanted to use it. This together with the way Usams handled VCADS Pro, and later, other applications, was the needed dynamicity that had to be offered to system responsible at the Volvo IT developing organisation.

- *A manager at Volvo truck says: It was our diagnostic tool called VCADS PRO that created a demand of a user-administration. And we needed a decentralised administration that controlled authorization and filter on each user. When Baldo was created, it was a tool that made us do a better business. We created a tool for Market companies to manage their users and their applications. Baldo is a simple solution on a complex problem. And Baldo turned out to be a success story.*

To do a better business is in high focus at a company as Volvo Trucks. Baldo gave Volvo Trucks an easy way of manage their users in a simple way. But there is a need of adjust application to fit a centralised user administration. This is hard to avoid without an enormous technically complex solution. Therefore Baldo is the way that is pointed out for today approximately 36000 user and 40 different applications.

7 SUMMARY

An Internet exposure will reach an enormous number of possible users. This is raising a number of new problems. Users must be administrated decentralised at a global market. Administrators at different continents can add links to a user's view of a portal. Many administrators, both central and local, create together with the user himself, a personalised outlook on the user's view of the portal. This situation is new for all IT-development and administration. Administrators can no longer know how the specific user's view of the portal looks like. The firewall is only accessed by the user and what links he has is only known by him. You can say that each user have their own "homepage". At Volvo AB's firewall it can be restrictions about what a user can add to his "homepage", but some freedom the user must be able to create a page he will benefit from.

New concepts as *Identity management* and *Access management* are in focus at the software industry. Different approaches to solve the *million user problem* on Internet are under development. But have a good enough administration tool being developed? To administrate hundreds of thousands users at the global market are a complex challenge. *The million user problem* can not be underrated in terms of man-time and economic costs.

The million user problem is a major task in this new environment. The problem can be described as an administrative problem. This enormous amount of users must be handled of administrators when they shall have maybe 10 to 100 applications each with right authorities, filters and actions. And those applications are placed in different environments such as UNIX, Windows, OS390, and AS400 and so on. Added to this are third part authority tools, used in those environments, how they work and how to administrate them. This is an impossible mission for a single administrator. Of course this problem is tried to be solved at different levels by many participants on the market. But have the recent development get long enough? A tool to handle authorities in all environments must be used. These reasons are the basis for the Baldo/Balda concept at Volvo AB. And for this, *the policy store* must use a more complex data model than can be offered by the common used standard LDAP.

Today companies use IT for their core business. More than 80% of all companies have their own portal and this percentage will rise even higher in the future. This gives an enormous market and many participants have today focus at *Identity management* and *Access management*. Some of the largest participants are giants such as IBM, Oracle and SAP. But the market show clear signals to be an early market. Signs as *Incomplete products*, *Poorly defined categories* and *Diverging visions*. Content managers need long term plans to handle their business but an early market can not give them a good decision base. Another aspect that companies on the Internet must consider is what the receiving media is. Today users on the Internet not only use PC. Also phones, PDA, low bandwidth, broadband and other hardware aspect must be handled. Dynamic solutions must be built, not only to handle different users, but also to handle different hardware.

Beside dynamic solutions for each user and hardware, we must consider *Cybertypes*. What different users look like and what race they belong to is not known on the Internet. Instead cultural differences to use Internet are to be considered. The use of symbols and what they mean in different cultures must be studied. Maybe shall colour and background change in different areas of the world to give a user a belonging to the portal? In the Western world humans and especially women are those who give most attention in advertisement, this is probably exactly the same in portals, but will that give same result in the Islamic world?

Three different solutions are compared and all have advantages and disadvantages.

The firewall solution gives high security and a possibility for all users on the same firewall to use information that are set in the cookie which is sent to all links that are connected. The technology that is used in most cases is a method called *early binding* and this technology offer a high level of security. This solution is the most common and it is used by the market leaders in this area. The disadvantage is that large amount of data is sent between http-headers when new links are activated. Another aspect is that no relational database is used. This gives a static and inflexible way to set user attributes. User administrators need more and flexible solutions for a complex activity. But the firewall solutions, especially the domains, in interaction with a relational database will make an administrative and security acceptable solution both for user administrators and security engineers. Sophisticated firewalls also offer domains or groups of links. *Early binding* domains are a really useful facility. To place a link in a domain gives you a firewall-controlled access that makes it impossible to access a link if you are not connected to the domain.

The database solution is a concept for a solution in one specific database. If you have control over all applications and are using the same database for all of them, it will be a good solution. A disadvantage is that all users need to be registered in the database. This must be done even if the user never or seldom accesses the portal. This will give a license cost for each registered user. Even database providers use LDAP as standard for user administration. As filters for presented information they often use triggers in the database. To make a database solution that interacts with solutions in other environments will be necessary as soon you have an application that does not exist in the same environment.

The Baldo/Balda solution can interact with LDAP, but have no need to do it. Instead Balda security engine hold all attributes for each user together with profiles, filters, granular access rights and accessible links, and it is saved in a relational database. The Balda *policy store* technique is called *late binding* and this offer a high level of dynamic solutions. At saves for specific links (or applications) a trigger sends user information to the LDAP directory (for the early binding), this is done to get a high level of security in the *early binding* technique. There are of course a lot of different needs in a company at Volvo's size, bought application, historical mainframe and so on. Therefore it necessary to have a dynamic solution and it must be accepted by other system managers. To set authorities for other applications in other environments need an acceptance inside the organisation.

One other major question is to create *single sign on* for a user at a portal. If a user has 20 applications, he does not want to logon every time he wants to start an application. This problem can be solved in all discussed solutions, but bought applications can not work exactly as you want it to, and you have not access to the source code. Therefore we need a combination of access methods to handle *single sign on*. The Baldo/Balda solution uses for e.g. *the firewall solution* to fetch logged on user from the cookie that is set at the firewall login.

Today most focus is to use the LDAP-standard as the solution of the administration problem. LDAP is a light version, developed to handle e-mails and attribute for e-mail users. For this purpose the LDAP-standard works very well, but can this be transformed to *the million user problem*? There is a discussion on the Internet about if LDAP is the right tool for companies that have a need to organize their users by *Identity management* concept. Can LDAP be complemented with missing functionality? I will say it can not be done, at least not without a combination of a module that takes care of the late binding. LDAP is a hierarchical solution without the flexibility a relational database can offer. But there are solutions that can interact with LDAP and that fill the gap between Internet demands and common standard. If this is done, we still benefit from a generally accepted industrial standard. At some point it is still necessary to exchange the LDAP-standard to a more Internet- suitable standard. When this point is reached is an object for discussion. Of course, LDAP can still be used for e-mail and other areas that it was once created for.

The Internet Engineering Task Force (IETF) is a large open international community. They discuss if LDAP should be used or not in email solutions that have different design goals than the X.500 standard had. The same questions must be asked if LDAP should be used to handle *Identity management* and *Access management* when it was not designed for that purpose.

What can LDAP do for you then? It can find things rapidly fast. Things as telephone number, sizes and colour. In a one-dimensional seeking it can find things that are connected to a single object. Directories are optimized for providing sophisticated searching capabilities of the data they hold. The LDAP standard is good at letting a large number of applications reach simple information.

One benefit when the *late binding policy store* is placed together with the portal is that we do not need to register the users in the database. The *late binding policy store* will control all buttons, fields, rights and roles at each application link.

How to solve the *million user problem* is a central question at the software industry today. Until more and dynamic solutions are standard many companies use LDAP for this purpose. But the LDAP standard is not designed for this problem. New standards must be considered. The Baldo/Balda at Volvo AB is a solution that can be a new way to look at the problem.

As an aspect I can say that my knowledge about the *Baldo/Balda solution* is far deeper than it is about *the firewall solution* and *the database solution*. But in this work I have tried to get a more circumstantial view of all three solutions.

A viewpoint is the problem to make those three solutions understandable for the respondents during the short interviews. A fact is that I had problem to make this different solutions understandable for any expert in such a short interview. Much more time and deeper discussions are relevant in a future work. But the general knowledge about the problem area of my respondents has been very good.

One question that must be considered is; if you are lifting out the user security from the database security module, and put it in the late binding, as in the Baldo/Balda solution at Volvo AB (figure 8), what consequences will this solution have? If you are placing the user security/authentication in a relation-database and connecting the database with the same database user for all managed users, can this be a solution that is accepted considering security reasons?

8 CONCLUSION

The key issue in my work have been a suggestion to describe a simple administration for a large amount of user, internally or at the internet or a combination of these. To solve this case there is a need of some kind of industry standard according to the interviews that are done.

This paper has also described the lack of this needed industry standard. But the demands are described and different approaches to solve the problems are done by different commercial companies. One of them is the Baldo/Balda solution at Volvo Information Technology.

Three main issues are identified to solve *the million user problem*.

The first one is *the release of local user administration* from local applications. To handle the administration for each user on different applications in different environments must be avoided. E.g. if administrators, that often are located globally, must know how to set authority in all possibly environment as Unix, OS390, AS400 and so on. It will be a nearly impossibly mission for the administrator to take care of hundreds of applications. But of course different applications have different needs, and a basket of solutions must be offered when you move identity management from the local application, to a central *late binding* module. Some applications can leave all their administration to an external tool. But other must get transactions that inform about changes on the user authority, and save it locally. Still the main issue is that this must be done in one single administration-tool and this is to make it simple for the administrators.

The second is distinguishing between *early and late binding*. Many solutions on the market make no difference between early and late binding. But my opinion is that this must be done, if we want to get an acceptable solution for a multiple application environment. We must let the early binding only take care of the user identification and create a high security to avoid unauthorized access to our systems. And the late binding gives instructions to any application "on-the-fly". And it is preferably that the late binding is handled by a relational database, to offer a dynamic and flexible solution for all different demands from different applications.

And the last one is the demand of *single-sign-on*. The frustration of non experienced users if they have to log on application locally must be avoided. This issue is maybe the most simple to solve, but it has also most impact on design of user-applications. If you put all http-applications on the same firewall you will get the possibility the move between different applications with the same log-on. The problem is that you are missing the functionality of control each user on your application. But this will be solved if you let the *late binding policy store* take care of this. And its here we need an industry standard, how to interact with an external module to get instructions for your application.

9 REFERENCES

Anupam, Vinpod. Breitbart, Yuri. Freire, Juliana. Kumar, Bharat. *Personalizing the Web Using Site Descriptions.* (1999). Bell Laboratories [online 2003-05-10] http://www.cse.ogi.edu/~juliana/pub/idm99.pdf

Backman, Jarl. (1998). Rapporter och uppsatser. Lund: Studentlitteratur

Baltimore Technologies. (2001). *Security Consideration for Progressive e-Corporations* [online 2003-05-10] http://www.baltimore.com/unicert/index.asp

Baltimore Technologies. (2002). *Enabling a Secure Business Platform* [online 2003-05-10] http://www.baltimore.com/unicert/index.asp

Björk, Lennart., & Räisänen, Christine. (1996). Academic Writing. Lund: Studentlitteratur

Dalton, John P., & Manning, Harly., & Gardiner Kathrine M. (2001). Forrester. *Managing Content Hypergrowth* [online 2003-05-10] http://www.vignette.com/Downloads/FR_MANAGING_CONTENT.pdf

Date C. J. (1990). Database Systems. Addison- Wessley

Fischer-Hübner, Simone. Quirchmayr, Gerald. Yngström, Louise. (1999). User identification & Privacy Protection. Sweden: DSV report series

Gebel Gerry (2003). Roles and Access Management: *Seeking a Balance Between Roles and Rules.* Burton Group. USA [online 2003-08-17] http://www.burtongroup.com

Gordon, B. Davis., & Olson Margarethe H. (1985). Management Information Systems. Minnesota: McGraw –Hill

Gunton, Tony. (1989). Infrastructure. UK: Prentice Hall

Haynes, John D. (2001). Internet Management Issues: A Global Perspective. USA: Idea Group Publishing

Hewitt, Joseph B. & Schumacher, Scott. & Weber, Gerald I. *Enterprise Identity Management Strategies* (2001). [online 2003-05-10] http://www.madison-info.com/Enterprise_Identity_Management.htm

Jacobsen, Jan Krag. (1993). Intervju. Lund: Studentlitteratur

Kahan José. (1995). *A capability-based authorization model for the World Wide Web.* Computer Networks and ISDN Systems 27 [online 2003-05-10] http://www.igd.fhg.de/archive/1995_www95/proceedings/papers/86/CAMWWW.html

Kahan José. (1999). *WDAI:a simple World Wide Web distributed authorization infrastructure.* Elsevier Science B.V [online 2003-05-10] http://www.elsevier.com/cas/tree/store/comnet/sub/1999/31/11-16/2153.pdf

Langefors Börje. (1995). Essays on Infology. Lund: Studentlitteratur

Lynch, Clifford. (1998). *A White paper on Authentication and Access Management Issues in Cross-organizational Use of Networked Information Resources.* Coalition for Networked Information [online 2003-05-10]
http://www.cni.org/projects/authentication/authentication-wp.html

Magoulas, Thanos., & Pessi Kalevi. (1991). En studie om informations-systemarkitekturer. Göteborg: Göteborgs Universitet, Department of Computer Sciences.

Magoulas, Thanos., & Pessi Kalevi. (1998). Strategic IT-management. Göteborg: Göteborgs Universitet, Department of Informatic.

Nakamura, Lisa. (2002). Cybertypes. Race, Ethnicity and Identity on the Internet. Great Britain: Routlege

Perry, Robert. (2001). *Managing the Content Explosion into Content-Rich Applications.* Yankee group [online 2003-05-10]
http://a144.g.akamai.net/7/144/812/fb9364709870f7/www.vignette.com/Downloads/AR_YANKEE_CONTENT.pdf

Poel van der, P.A.M.M. (1989). Framework for Architectures in Information Planning, in International Operations, Crossing Borders in Manufacturing and Service. Elisvier Science Publishers.

Sommerville, Ian. (1989). Software Engineering. England: Addison-Wesley

Tandenbaum, S. Andrew. (1988). Computer Networks. The Netherlands: Prentice –Hall

The Fact Point group (2003). *Accessing the Economic Benefits of Oracle9iAS Portal*. USA [online 2003-05-17]
http://www.oracle.com/ip/deploy/ias/portal/Oracle9iASPortal_FactPointGroup.pdf

The Internet Engineering Task Force (2003). *Why not use LDAP?.* USA [online 2003-05-10]
http://www.ietf.org/proceedings/96dec/app/acap-slides/tsld005.htm

Tim Howes, Mark C. Smith, Gordon S. Good, Timothy A. Howes. (1998). Understanding and Deploying LDAP Directory Services. UK: Sams Publishing.

University of Maryland Biotechnical Institute (2003). *Other-email.* USA [online 2003-05-10]
http://www.umbi.umd.edu/computing/netscape/why_netscape.html

Uwe, Jendricke. Kreutzer, Michael. Zugenmaier, Alf. [A] (1997). *Mobile Identity Management*.Germany: Albert-Ludwigs-University of Freiburg [online 2003-05-10]
http://citeseer.nj.nec.com/543199.html

Uwe, Jendricke. Kreutzer, Michael. Zugenmaier, Alf. [B] (1997). *Pervasive Privacy with Identity Management*.Germany: Albert-Ludwigs-University of Freiburg [online 2003-05-10]
http://citeseer.nj.nec.com/544380.html

Wainewright, Phil (2002). *Web Services Infrastructure.* Loosely Coupled USA [online 2003-11-02]
http://www.philwainewright.com/pubs/wp/WSIpaper.pdf

Wallström, Martin. (2003). Oracle leder på portaler. *Computer Sweden nr 19*

Yasin, Rutrell. (2002). *What is identity management?,* TruSecure Corporation White paper [online 2003-05-10]. http://www.infosecuritymag.com/2002/apr/cover_casestudy.shtml

9.1 Figures and Tables

10 APPENDIX

10.1 Codd's rules

10.1.1 The foundation rule

This is Codd's Rule Zero which is to foundation for the other rules. The rules states that:

Any system which claims to be a relation database management system must be able to manage databases entirely through its relational capabilities
and means that the RDBMS must support:

A data definition language
A data manipulation language
A data integrity language
A data control language

All of which must work on several records at a time, that is, on a relation.

10.1.2 Information rule

All information in a RDB must be represented explicitly at the logical level in just one way, that is, by values in tables.

10.1.3 Guaranteed access rule

Each item of data in a RDB must be guaranteed to be logically accessible by using a combination of table name, primary key name, and primary key value.

10.1.4 Systematic nulls rule

Null values are distinct from an empty character string, a string of blank characters, a zero or any other number, and must be supported by representing missing and/or inapplicable information in a systematic way.

10.1.5 Dynamic catalogue rule

The description of the DB must be represented at the logical level just like ordinary data. Authorised users must be able to use the data manipulation language to interrogate the DB in the same way that they interrogate ordinary data.

10.1.6 Comprehensive data sub-language rule

A RDB must support at least one language which is capable of supporting:

Data definition
View definition
Data manipulation
Interactively
By program in a conventional host language
Integrity constraints
Authorisation
Transaction boundaries

10.1.7 View updating rule

All views of the data which are theoretically updatable must be updatable in practice by the DBMS.

10.1.8 High-level language rule

The capability of handling a relation as a single operation applies to:

The retrieval of data
And also:

The insertion of data
The updating of data
The deletion of data

And these must be possible:

Interactively By program in a conventional host language

10.1.9 Physical data independence rule

Application systems and terminal activities must be logically unimpaired whenever any changes are made to the way in which the physical data is stored, the storage structures which are used, or the access methods.

10.1.10 Logical data independence rule

Application systems and terminal activities must be logically unimpaired whenever any changes are made to the way in which the logical data is organised, such as when a table is dispersed to several tables to optimised storage or performance.

10.1.11 Integrity independence rule

Integrity constraints must be definable in the RDB sub-language and stored in the system catalogue and not within individual application programs.

10.1.12 Distribution independence rule

Application systems and terminal activities must be logically unimpaired whenever the data is redistributed amongst several locations on a data communications network.

10.1.13 Non-subversion rule

If the DB has any means of handling a single record at a time, that low-level of working must not be able to subvert or avoid the integrity rules which are expressed in a higher-level language which handles multiple records at a time

10.2 Interview questions outside Volvo

10.2.1 Can you describe your work position at the Company?

10.2.2 Can you describe your experience of Identity and Access management?

10.2.3 What strategy does your company use for handling questions about Identity and Access management?

10.2.4 What standards do your company use, according to Identity and Access management?

10.2.5 What advantages/disadvantages can you see with the standards today?

10.2.6 What differences do you see between traditional user administration and Identity management?

10.2.7 Do you see any cultural differences between different countries in the way of using IT and portals?

10.2.8 Can you see different behaviour of users in different countries?

10.2.9 Can you find the competence and tools you need at the market?

10.2.10 Do you see any problems or possibilities in the future?

10.3 Interview questions inside the Volvo organisation

10.3.1 What influence did Baldo had on the Volvo Parts organisation?

10.3.2 What influence did Baldo had on the organisation at Volvo Partners?

10.3.3 What experience do users have concerning web-solutions?

10.3.4 Does Partners to Volvo have enough infrastructures?

10.3.5 What is Baldo’s biggest advantage?

10.3.6 Is there any disadvantage with Baldo?

10.3.7 Why is Baldo a success story?