Identity and Access Control - demonstrating compliance

Marc Sel

Bart Van Rompay

PricewaterhouseCoopers

marc.sel@pwc.be, bart.van.rompay@pwc.be

Abstract

Identity and particularly access control present various challenges, particularly for larger organisations. The combined complexity of users from various communities, accessing multiple systems and applications in the context of business processes can be significant. The US NIST proposed the Role-Based Access Control model in order to effectively and efficiently manage authorisations. While this model certainly also has its drawbacks, it gave rise to various interesting software solutions. One particularly relevant one is the Sage tool. This tool builds a model of the actual authorisations across platforms by consolidating and enriching them in its own database. Subsequently, the built-in pattern-matching engine can identify a number of less desirable patterns in the data and can recommend solutions, e.g., for role structuring (role-mining). Furthermore, business constraints can be expressed in so-called business process rules, which can, e.g., reflect segregation of duty requirements.

In the pilot project described here as case study, we combined both role-mining and compliance verification. The case study organisation is subject to both national competition regulation and the US Sarbanes-Oxley act. They employ approximately 25.000 employees. Analysing existing access controls through a unified approach and applying compliance rules to them has shown to be a quick and reliable way for them to demonstrate compliance (or identify actions where compliance was not yet achieved). The fact that the control library is available both at the level of principles and at the level of specific business process rules makes the approach transparent, repeatable and affordable. Furthermore a number of observations were made that allowed to remove undesired authorisations through data cleaning. As a result of the pilot project the client decided to implement BPR-based compliance verification for all applications that are subject to Sarbanes-Oxley.

1 The challenge of Identity and Access Control

1.1 Introduction

Most medium to large sized organisations today built up and manage what could be referred to as their ‘authorisation space’. This space is essentially structured into three dimensions: the different user communities (subjects), the ICT services and applications (objects), and the processes allocating users authorisations onto these services.

This can be represented as:

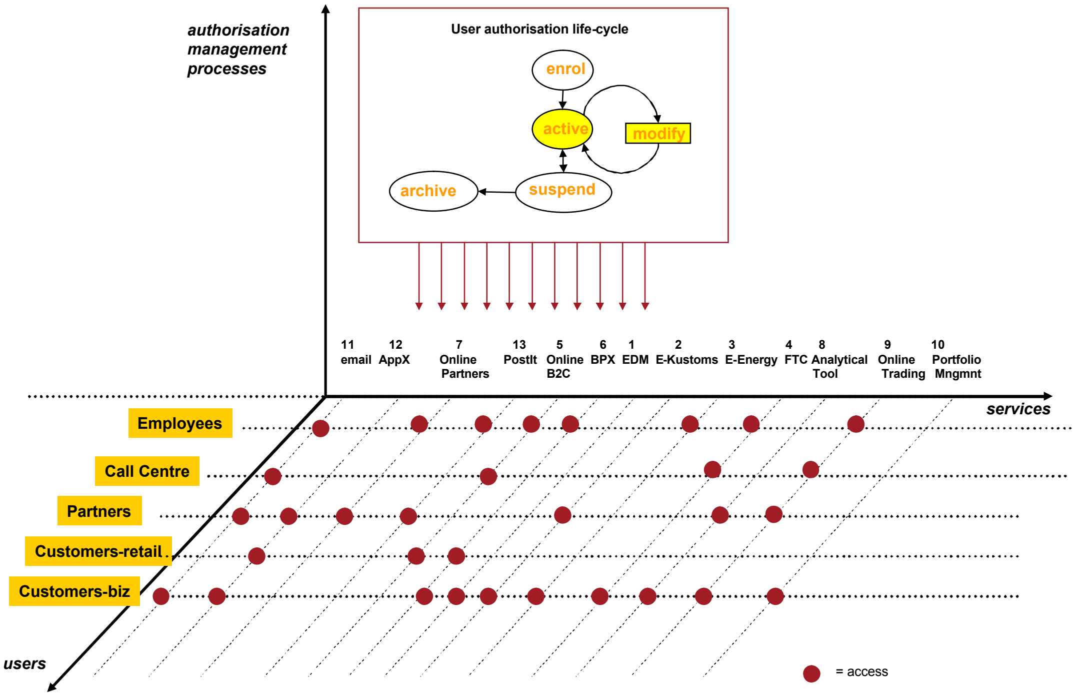


Figure 1: The three dimensional authorisation space

The user dimension (subjects) is structured into various types of user communities, ranging from employees to partners and customers, or the public at large. Nowadays, even some regulators are asking access, or are forcing companies to open systems to competitors in order to liberalise a particular market. The services dimension (objects, left-to-right axis) can be further decomposed into individual applications or transactions within these applications. Finally, the authorisation management dimension (vertical axis) is organised into sets of processes that support the user authorisation life-cycle (“from hiring to firing/retirement”).

In the real-world, this space can be impressively large. For one particular company with 40.000 employees (and excluding the authorisations of customers on company systems) we estimated the total number of authorisations that were managed around 35 million. Since that organisation’s authorisations were decided by a core team of 10 persons, this meant that on average, every authorisation manager was dealing with approximately 3,5 million authorisations. Most of these authorisations have been built up over the years, often surviving multiple rounds of business reorganisation.

1.2 IdM initiatives often fall short of meeting expectations

Many vendors tout Identity Management (IdM) systems as the overarching solution to the management of user identification and authorisation. Such systems are aiming essentially at quicker turnaround time for user-id and authorisation provisioning. These systems typically address the aspects of authentication, directories, provisioning and access control. While the actual success rate of such Identity Management projects varies, their approach with regard to access control is typically incomplete. High-level or coarse-grained access control can be managed, but more fine-grained or application-specific access control is often not addressed. Also, while web-based solutions are typically covered, legacy systems are often left out. For many organisations legacy systems will stay around for the near or not-so-near future.

In practise, such projects are often delayed or less successful than expected due to the overwhelming combinatory complexity. An organisation with ten-thousands of users and hundred-thousands of resources (applications, databases, files, …) quickly has multi-millions of authorisations to manage across both legacy, ERP and web systems. Automating the existing authorisations in a new IdM system may fall prey to the old adagio “garbage in – garbage out”.

Furthermore, while we think that quicker turn-around time and improved user and authorisation management are valid objectives, these are still missing an important point. Equal attention should be paid to understanding the structure of the authorisation space in order to reduce and restructure it in order to facilitate more effective and efficient management processes over the authorisations.

Adding to the above, there is a clear increase in regulation resulting in ever more complex compliance requirements. Most organisations have to deal with regulation at three levels at least:

- Global – e.g., directives such as various directives with regard to privacy, anti-money laundering, electronic signatures etc;
- Industry-sector specific – e.g., with regard to market regulation and/or liberalisation. Examples include the US-originated Sarbanes-Oxley act or an EU national law on Competition;
- ICT specific – e.g., ISO/IEC 27001 or 17799 for information security.

So most organisations find themselves confronted with both a complex authorisation space to manage and the requirement to do this in a sufficiently transparent and understandable way. The onus of demonstrating this is imposed on the organisation.

2 The way forward

2.1 Increasing abstraction

In order to effectively and efficiently manage such a complex space, a high degree of abstraction is mandatory. Traditional mechanisms such as Bell-Lapadula and the access control matrix have evolved into today's model of roles. Adequate access control management requires abstraction to make decisions, while the business processes demand sophisticated and often complex IT systems and infrastructures. These IT systems typically include authorisation mechanisms and repositories from many different backgrounds and technologies. However, good management (and extending it to governance and compliance) requires a unified approach to address the many possible constraints that have to be taken into account, such as location-based limitations, confidentiality, segregation-of-duty, functional limitations etc. We are convinced that the most effective way forward is to increase the level of formalisation by modelling the complex multi-technology access space.

2.2 A possible way forward

We see the most realistic way forward as the combination of two key elements. On one hand: unifying technologies, and on the other hand, control libraries.

2.2.1 Unifying technologies

Mathematics came to the rescue by allowing us to formulate models such as RBAC – role based access control (which originated from the US NIST – [NIST2001]). This is an overarching model for authorisation management that is gradually gaining acceptance and standardisation.

The basic model is based on a "User-Role-Permission" paradigm. Since users typically access a system via (multiple) sessions, this is also reflected.

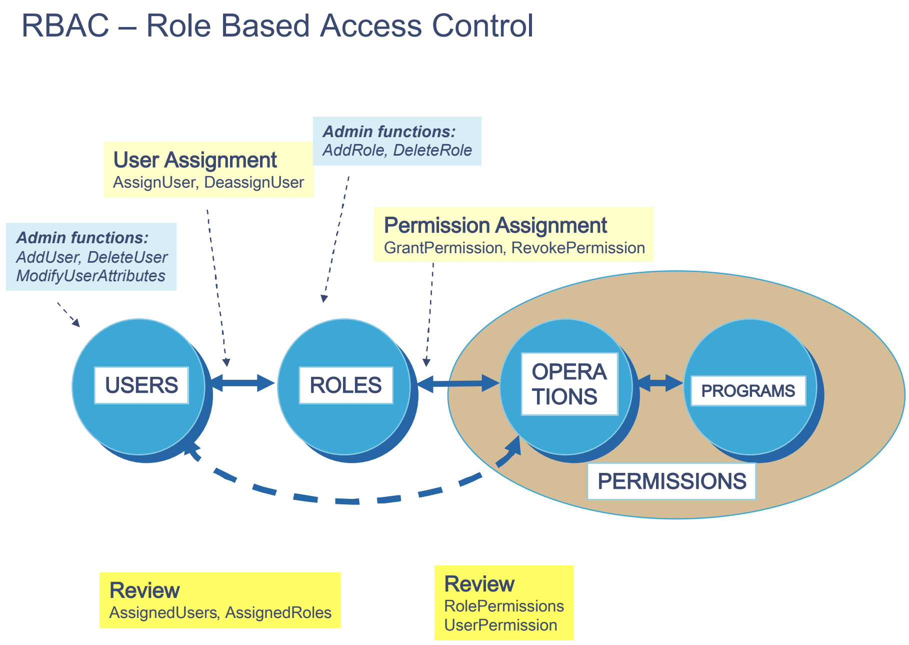

RBAC – Role Based Access Control
Admin functions: AddRole, DeleteRole
User Assignment
AssignUser, DeassignUser
Admin functions: AddUser, DeleteUser ModifyUserAttributes
Permission Assignment
GrantPermission, RevokePermission
USERS
ROLES
OPERA TIONS
PROGRAMS
PERMISSIONS
Review
AssignedUsers, AssignedRoles
Review
RolePermissions
UserPermission

Figure 2: RBAC at a glance

The role model was widely accepted, and is currently still gaining popularity. It allows to model the diverse authorisations in an organisation quite well, striking a balance between effectiveness and efficiency. Nevertheless, we want to mention two areas where we think the model falls somewhat short for meeting the requirements of larger-scale commercial organisations:

- Within such organisations, there is often a natural grouping of users into user communities on one hand and permissions (or resources) into permission communities on the other hand. This leads to a four tier-model that is used in practice (user – user cluster – permission cluster – permissions). However, it is not easily modelled in RBAC. It can be argued that a role-hierarchy makes this possible, but this is rather theoretical in our opinion. Larger-scale commercial organisations often prefer to have more conceptual layers to model their authorisations than just three.
- Also, within such organisations, the organisational dimension itself is both very important and changing at a high pace (e.g. due to mergers and acquisitions). Larger-scale commercial organisations often have a need to model the organisation itself in their authorisations model.

RBAC's main competitor is probably the rule-based model, where access is dynamically calculated by rules that evaluate the values of attributes (e.g. LDAP attributes). We will not discuss this rule-based access control model here.

Based on the role model, additional models and algorithms were introduced that are able to analyse the actual patterns exposed by the existing authorisation repositories. Pattern recognition algorithms can quickly sift through vast amounts of data and identify patterns that are out-of-sync with normal expectations, or can propose improved structures. Provisioning mistakes (the account receivable clerk with system-level access) or redundant role definitions can be immediately identified (and rectified).

Furthermore additional formalisms can be defined that specify constraints (e.g. segregation-of-duty) on the data. Applying those formalised constraints allows testing for compliance.

2.2.2 Control library

We use the name control library to refer to a collection of control principles. By control principles we refer to the typical controls defined by an organisation in the context of internal control, or a preparation for Sarbanes-Oxley compliance, or comparable. The first COSO report can probably be considered as the seminal paper in this regard. While regulation or COSO typically outline the requirements for a. o. internal controls, they remain at a high and often abstract level. In practise, this abstract level is translated into more pragmatic sets of controls, that we will refer to as a control library.

Such a control library typically contains definitions that address:

- restriction on organisational scope (coarse grained/fine grained access, across organizational and/or legal entities)
- access to critical transactions
- segregation of duty (SOD)
- orphans (i.e., entities such as users, roles or resources present in the system but not connected to another tier. As such they contribute to the complexity but not to the actual access control definitions)
- collectors (people or processes that accumulated a significant amount of authorisations, typically over a longer period of time and due to a lack of good management processes).

We propose a three-tier structure for a control library that is focused on identity and access management:

- tier #1: the control baselines;
- tier #2: controls related to organisational structure and processes;
- tier #3: controls related to time.

The control baselines (tier #1) specify commonly accepted principles with regard to identity and access management, such as

- Individual accountability – authorisations are granted to specific individual users (which can be physical persons or technical users such as daemons or address spaces). As a consequence, userids/accounts are not shared;
- Single user identification – a user should have a single identifier per platform. Situations where users have different identifiers across platforms or multiple identifiers on the same platform should be avoided;
- Authorisations should be allocated through roles (or a similar grouping mechanism). Direct links between users and resources should be avoided;
- No single user should have all authorisations. If such users need to exist from a technical perspective, they should be blocked for daily operational activities;

- No users should accumulate so many authorisations that there can be reasonable suspicion that the risk for (un-)intentional misbehaviour increases. There should be a form of monitoring that users should not accumulate a set of authorisations that deviates significantly from their peers or comparable functions (unless justified).
- There should be no "orphans" in the identity and access management system, i.e., there should be no users without authorisations connected to them, there should be no roles without users and authorisations connected to them, and there should be no authorisations without users connected to them[1].

The controls related to organisational structure and processes (tier #2) should reflect that a user performs a function (role) within an organisation, typically in the context of one or more processes. This leads to the following principles:

- Authorisations should be limited to the appropriate functional organisational scope and processes. Where required this may lead to 'Chinese Walls (or the well-known Brewer-Nash model)';
- Authorisations should reflect a high-level segregation between production, acceptance/test and development environments;
- Authorisations should reflect the required segregation-of-duties (combinations of certain authorisations are to be forbidden);
- Specific functions within the organisation require specific authorisations. For example, auditors will have read authorisations only.

The controls related to time (tier #3) should reflect the fact that only active users and active authorisations need to be present in the system:

- Users that are no longer employed or servicing the organisation need to be blocked;
- Users that have not accessed the systems for the last 90 days need to be blocked

Please note that this is an example of a typical implementation of a control library for identity and access management. It should not be considered as automatically suitable for a particular environment, but rather it should be validated.

2.2.3 Combining unification and control libraries

Organisations rely on many applications across multiple platforms (SAP, Oracle, in-house developments, …). There are many solutions available today that address the definition of controls as well as their enforcement and compliance. Most identity management solutions cover

- [1] Obviously the organisation may keep expired users and authorisations for historical reasons, these should however be separated from the active set.

parts of this spectrum. Approva and Virsa focus specifically on the controls aspects. We used the Sage tool from Eurekify[2] in the case study. This tool allows unifying the authorisation data from the different platforms, and allows cleaning and optimising this authorisation data. For information on the mathematical foundation of the tool, refer to [Rymon93]. Furthermore, it allows defining so-called Business Process Rules, which can be used to test for compliance.

3 Case study

3.1 The challenge

The case study organisation is a European company that is subject to both national competition regulation and the US Sarbanes-Oxley act. They employ approximately 25.000 employees. The company will be subject to Sarbanes-Oxley compliance audits as from January 2007.

They recognised the need to strictly manage authorisations, and initiated a company-wide identity management project. It is expected that this project will deliver what is required from an authorisation management perspective in due time. However, demonstrating regulatory compliance is not intended to be a direct outcome of the project. The primary challenge was to define compliance rules and to demonstrate that the actual authorisations across the various applications and systems comply with those. There were approximately 60 applications identified for which compliance needs to be periodically demonstrated. We will now discuss the case of one specific application, which we will call PICASSO.

3.2 The solution

A team from PwC was invited to create a solution. We based our solution on the application of our in-house developed control libraries on the PICASSO application, and on the Sage tool from the company Eurekify. We loaded the user and authorisation information, analysed the structure and contents of the authorisations and defined compliance rules. Analysing the structure and contents of the authorisations allowed us to recommend changes to the actual authorisations in place. Applying the compliance rules to the authorisation data allowed us to identify compliance violations (or through the lack of these, demonstrate compliance).

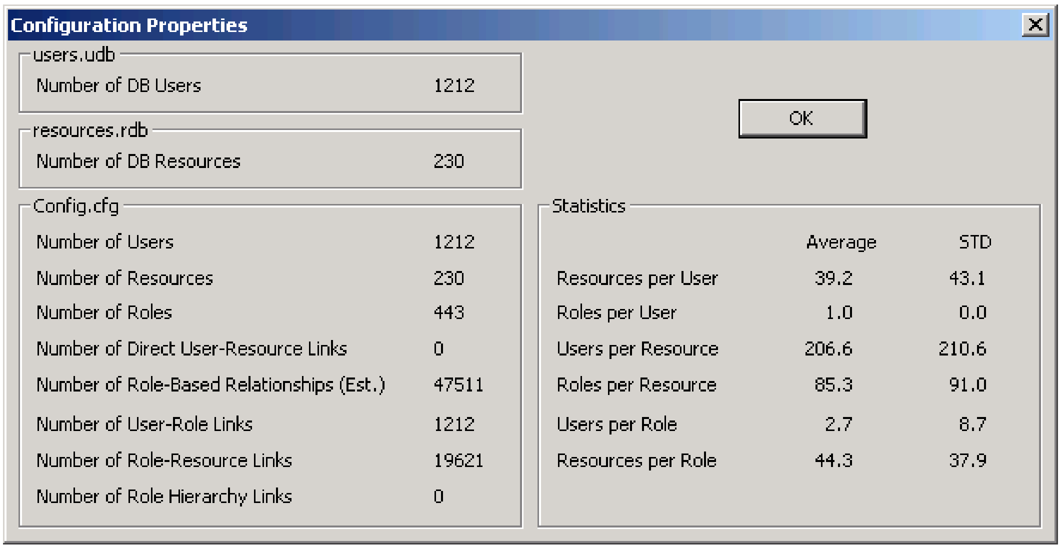
3.3 Role-mining the authorisation data

Initially we used the Sage product to identify potential improvements in the structure and contents of the authorisation data. This information was then fed back to the identity management project and the application owners.

[2] www.eurekify.com

The following tasks were performed. The authorisation files were extracted from the PICASSO application. This data was loaded into Sage via CSV[3] files, and enriched with user attributes from the HR database.

We then had the following basic authorisation data available:



users.udb

Number of DB Users	1212

resources.rdb

Number of DB Resources	230

Config.cfg

Number of Users	1212
Number of Resources	230
Number of Roles	443
Number of Direct User-Resource Links	0
Number of Role-Based Relationships (Est.)	47511
Number of User-Role Links	1212
Number of Role-Resource Links	19621
Number of Role Hierarchy Links	0

Statistics

	Average	STD
Resources per User	39.2	43.1
Roles per User	1.0	0.0
Users per Resource	206.6	210.6
Roles per Resource	85.3	91.0
Users per Role	2.7	8.7
Resources per Role	44.3	37.9

Figure 3: The dimensions of the actual PICASSO authorisations – Sage screen capture

As one can easily see, this configuration handled the authorisations of 1.212 users, via 443 roles onto 230 resources. There were no direct links from users to resources (as dictated by 'best-practice').

We made the following observations:

- 5 roles (32 users) have all resources – this is not in line with good practice;
- 22 users had no access to any resources at all – they were only present for historical reasons;
- 251 of 443 roles have no users at all (due to reorganizations – should be cleaned on a short term);
- 74 roles have only 1 user;
- Many sets of roles exist with the same (or almost the same) resources.

Furthermore, a significant number of users could not be related to the official HR database.

These observations were communicated to the Identity Management project which could take them into account when defining and simplifying their authorisation infrastructure.

However, it became quickly apparent that the compliance problem was even more important, due to the fact that a lack of compliance immediately results in business problems such as violating the US Sarbox policy, which has an immediate and significant impact (such as loosing the right to be quoted on a US stock exchange).

[3] CSV: Comma Separated Values

3.4 Defining and testing compliance via business rules

We implemented compliance rules via the business process rules of Eurekify's Sage tool. We will now discuss the consecutive steps performed and results obtained. These steps are: the definition of compliance drivers, the specification of business process rules, and the testing of the authorisation data according to those rules.

The first step consists in identifying the compliance drivers. We identified:

- The existing authorisations matrix, manually maintained in Excel;
- Restriction of a particular resource (PICASSO function) to specific employee classes - access to function F5909 restricted to billing employees (role R-HSE-BLL) and TNU disturbance analysts (role R-BPX089);
- Restriction of a particular function combination to a specific employee class - access to the combination of functions F5909-F5326 restricted to billing employees (role R-HSE-BLL);
- Users belonging to the 'retail' organisational unit may only have 'read' access.

Once all compliance drivers are identified, they are translated into BPR's (business process rules). These are XML files that specify constraints over the various elements of the authorisation data. The same authorisation data that were loaded for the role-mining can be reused. There are three types of BPR's that can be specified. They are referred to as business constraints, segregation of duty, and license.

The first type, business constraints, allow to express constraints on the following combinations:

- Role-Role – a restriction on the users in two sets of roles;
- Role-Resource – a restriction between the users in a set of roles and a set of resources;
- Resource-Resource – a restriction on the users in two sets of resources;
- User Attribute – Role – a restriction between users with a certain attribute value and a set of roles;
- User Attribute – Resource – a restriction between users with a certain attribute value and a set of resources.

The constraints that can actually be expressed are:

- Forbidden – Users in left side are not allowed to be on right side
- Must be – Users in left side must also be on right side
- Only allowed – Users in left side are only allowed to roles/resources on right side

- May be – Only users in left side (and not others) are allowed to roles/resources on right side

The second type, „segregation of duty“ allows to express segregation of duty constraints either at the level of roles or at the level of resources. Finally, the third type is oriented towards license verification. As this was out of scope for the project we will not elaborate on it any further.

Let us now provide an example. The second compliance driver (access to function F5909 is restricted to billing employees (role R-HSE-BLL) and TNU disturbance analysts (role R-BPX089) is expressed as the following BPR-rule:

```
<BPR>
      <ENTRY TYPE="1201" ID="F5909" DESCRIPTION="Function 5909 should be
      restricted to billing functions and SNT disturbance analysts">
            <LEFT F1="R-HSE-BLL"/>
            <LEFT F1="R-RPX069"/>
            <RIGHT F1="RELAY BDAF" F2="5909" F3="380"/>
      </ENTRY>
</BPR>
```

Obviously, the compliance analyst is not required to manually encode XML statements, the BPR definitions are created through a GUI.

Finally, testing the BPR’s over the actual PICASSO authorisation data led to the following observations.

With regard to the first compliance driver (authorisations matrix), we identified six violations in the actual authorisation data. Both the second (functional restriction) and third (restriction on functional combination) compliance drivers resulted in six violations each. With regard to the fourth compliance driver (only ‘read’ access for members of the retail organisation) we found in total 151 violations. This included three persons that were within retail but simply had all possible authorisations.

4 Conclusion

Analysing existing access controls through a unified approach and applying compliance rules to them has shown to be a quick and reliable way for this particular organisation to demonstrate compliance (or identify actions where compliance was not yet achieved).

The fact that the control library is now available both at the level of principles (based on Sarbanes-Oxley and other regulations) and at the level of specific business process rules makes the approach both transparent and repeatable.

Furthermore, a number of observations were made that allowed to remove undesired authorisations through data cleaning.

As a result of the pilot the client decided to implement BPR-based compliance verification for all applications that are subject to Sarbanes-Oxley.

References

[Rymon93] Rymon, Ron: An SE-tree based Characterization of the Induction Problem. In: Proceedings Machine Learning Conference, Amherst, MA, 1993.

[NIST2001] ACM Transactions on Information and System Security, Vol. 4, No. 3, August 2001, pages 224-274.

Keywords

Compliance

Identity Management

Role-mining

RBAC

Role Based Access Control

Sarbanes-Oxley

Segregation of duty