Refactoring Access Control Policies for Performance Improvement

Donia El Kateb
Laboratory of Advanced Software SYstems (LASSY)
University of Luxembourg
Luxembourg
donia.elkateb@uni.lu

Tejeddine Mouelhi
Security, Reliability and Trust Interdisciplinary Research Center, SnT
University of Luxembourg
Luxembourg
tejeddine.mouelhi@uni.lu

Yves Le Traon
Laboratory of Advanced Software SYstems (LASSY) &
Security, Reliability and Trust Interdisciplinary Research Center, SnT
University of Luxembourg
Luxembourg
yves.letraon@uni.lu

JeeHyun Hwang
Dept. of Computer Science
North Carolina State University
U.S.A
jhwang4@ncsu.edu

Tao Xie
Dept. of Computer Science
North Carolina State University
U.S.A
xie@csc.ncsu.edu

ABSTRACT

In order to facilitate managing authorization, access control architectures are designed to separate the business logic from an access control policy. To determine whether a user can access which resources, a request is formulated from a component, called a Policy Enforcement Point (PEP) located in application code. Given a request, a Policy Decision Point (PDP) evaluates the request against an access control policy and returns its access decision (i.e., permit or deny) to the PEP. With the growth of sensitive information for protection in an application, an access control policy consists of a larger number of rules, which often cause a performance bottleneck. To address this issue, we propose to refactor access control policies for performance improvement by splitting a policy (handled by a single PDP) into its corresponding multiple policies with a smaller number of rules (handled by multiple PDPs). We define seven attribute-set-based splitting criteria to facilitate splitting a policy. We have conducted an evaluation on three subjects of real-life Java systems, each of which interacts with access control policies. Our evaluation results show that (1) our approach preserves the initial architectural model in terms of interaction between the business logic and its corresponding rules in a policy, and (2) our approach enables to substantially reduce request evaluation time for most splitting criteria.

Categories and Subject Descriptors

D.4.8 [**Performance**]: Measurements; C.4 [**Performance of systems**]: Performance attributes

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ICPE'12, April 22-25, 2012, Boston, Massachusetts, USA
Copyright 2012 ACM 978-1-4503-1202-8/12/04 ...$10.00.

General Terms

Performance, Design

Keywords

Access Control, Performance, Refactoring, Policy Enforcement Point, Policy Decision Point, eXtensible Access Control Markup Language

1. INTRODUCTION

Access control mechanisms regulate which users could perform which actions on what system resources based on access control policies. Access control policies (policies in short) are based on various access control models such as Role-Based Access Control (RBAC) [7], Mandatory Access Control (MAC) [6], Discretionary Access Control (DAC) [10], and Organization-Based Access Control (OrBAC) [9]. Access control policies are specified in various policy specification languages such as the eXtensible Access Control Markup Language (XACML) [3] and Enterprise Privacy Authorization Language (EPAL) [1]. A policy-based system allows policy authors to define rules that specify actions (e.g., read) that subjects (e.g., students) can take on resources (e.g., grades) in a policy. In the context of policy-based systems, an access control architecture is often designed with respect to a popular architectural concept that separates Policy Enforcement Points (PEPs) from a Policy Decision Point (PDP) [19]. More specifically, a PEP is located inside an application's code (i.e., business logic of the system). Business logic describes functional algorithms to govern information exchange between access control decision logic and a user interface (i.e., presentation). Given requests (e.g., student A requests to read her grade resource B) formulated by the PEP, the PDP evaluates the requests and returns their responses (e.g., permit or deny) by evaluating these requests against rules in a policy.

An important benefit of such architecture is to facilitate managing access rights in a fine-grained way by decoupling the business logic from the access control decision logic, which can be standardized and separately managed. However, this architecture may cause performance degradation especially when policy authors maintain

a single policy with a large number of rules to regulate the whole system's resources. Consider that the policy is centralized with *only* one single PDP. The PDP evaluates requests (issued by PEPs) against the large number of rules in the policy in real-time. Such centralization can be a major factor for degrading performance as our previous work [13] showed that efficient request evaluation with a large number of rules is a challenging task. This performance bottleneck issue may impact service availability as well, especially when dealing with a huge number of requests within a short time.

In order to address this performance bottleneck issue, we propose an approach to refactor policies automatically to significantly reduce request evaluation time. As manual refactoring is tedious and error-prone, an important benefit of our automated approach is to reduce significant human efforts as well as improving performance. Our approach includes two techniques: (1) refactoring a policy (handled by single PDP) to its corresponding multiple policies each with a smaller number of rules (handled by multiple PDPs), and (2) preserving the architectural property stating that a single PDP is triggered by a given PEP at a time.

In the first technique, our approach takes a splitting criterion and an original global policy (i.e., a policy governing all of access rights in the system) as an input, and returns a set of corresponding sub-policies, each of which consists of a smaller number of rules. This refactoring involves grouping rules in the global policy into several subsets based on the splitting criterion. More specifically, we propose a set of splitting criteria to refactor the global policy to smaller policies. A splitting criterion selects and groups the rules handled by the overall PDP into specific PDPs. Each criterion-specific PDP encapsulates a sub-policy that represents a set of rules that share the same combination of attribute elements (Subject, Action, and/or Resource). In the second technique, our approach aims at preserving the architectural property that only a single PDP is triggered by a given PEP at a time. More specifically, given a request, each PEP should be mapped to a PDP loaded with a policy, which includes a set of rules to be applicable for the request. Therefore, our refactoring maintains the architectural property of centralized architectures in policy-based systems.

We collect three subjects of real-life Java systems. Each system interacts with access control policies, whose corresponding request evaluation faces performance degradation. The policies are specified in eXtensible Access Control Markup Language (XACML) [3]. XACML is an XML-based policy specification language popularly used for web-based applications and services.

While our subjects are based on XACML policies, our approach could be applicable to any software system that interacts with policies specified in other policy specification languages. We conduct an evaluation to show performance improvement achieved by our approach in terms of request evaluation time. We leverage two types of PDPs to measure request evaluation time. The first one is the Sun PDP implementation [2], which is a popular open source PDP, and the second one is XEngine [13], which transforms an original policy into its corresponding policy in a tree format by mapping attribute values with numerical values. Our evaluation results show that our approach preserves the policy behaviors of the centralized architectures and the architectural property. Our evaluation results also show that our approach enables reducing the request evaluation time substantially. This paper makes the following three main contributions:

- We propose an automated approach that refactors a single global policy to policies each with a smaller number of rules. This refactoring helps improve performance of request evaluation time.
- We propose a set of splitting criteria to help refactor a policy in a systematic way. Our proposed splitting criteria do not alter policy behaviors of the centralized architectures.
- We conduct an evaluation on three Java systems interacting with XACML policies. We measure performance in terms of request evaluation time. Our evaluation results show that our approach achieves substantially faster than that of the centralized architectures in terms of request evaluation time.

The remainder of this paper is organized as follows. Section 2 introduces concepts related to our research problem addressed in this paper. Section 3 presents the overall approach. Section 4 presents evaluation results and discusses the effectiveness of our approach. Section 5 discusses related work. Section 6 concludes this paper and discusses future research directions.

2. CONTEXT/PROBLEM STATEMENT

This section further details a centralized architecture, its two desirable features such as synergy and reconfigurability, and its induced penalty (performance bottlenecks). Managing access control policies is one of the most challenging issues faced by an organization due to frequent changes in a policy. For example, a policy-based system has to handle some specific requirements such as role swapping when employees are given temporary assignments, as well as changes in the policies and procedures, new assets, users and job positions in the organization.

2.1 Centralization of Architectures

To facilitate policy management, an access control policy is traditionally modeled, analyzed, and implemented as a separate component encapsulated in a PDP. This separation leads to the centralized architecture presented in Figure 1, in which one single PDP is responsible for granting/denying the accesses that are requested. This centralized architecture is a simple solution to easily handle changes in policy-based systems by enabling the policy author to directly change policies on the single PDP. The separation between the PEP and the PDP simplifies policy management across many heterogeneous systems and limits potential risks arising from incorrect policy implementation or maintenance when the policy is hardcoded inside the business logic.

2.2 Centralization: A Threat for Performance

In such a centralized system, when a service regulated by an access control policy requires an access to some resources in the system, the PEP calls the PDP to retrieve an authorization decision based on the policy encapsulated in the PDP. This authorization decision is made through the evaluation of rules in the policy. Subsequently, an authorization decision (permit/deny) is returned to the PEP. When a huge number of access requests are sent by the PEP to the PDP, two bottlenecks cause performance degradation:

- all the access requests have to be managed through the same input channel of the PDP.
- the centralized PDP computes an access request by searching which rule is applicable among all the rules that the encapsulated policy contains.

A request evaluation time is thus strongly related to

- the number of rules in the policy that the PDP contains [14].
- the workload (i.e., the number of requests) that have to be evaluated by the system.

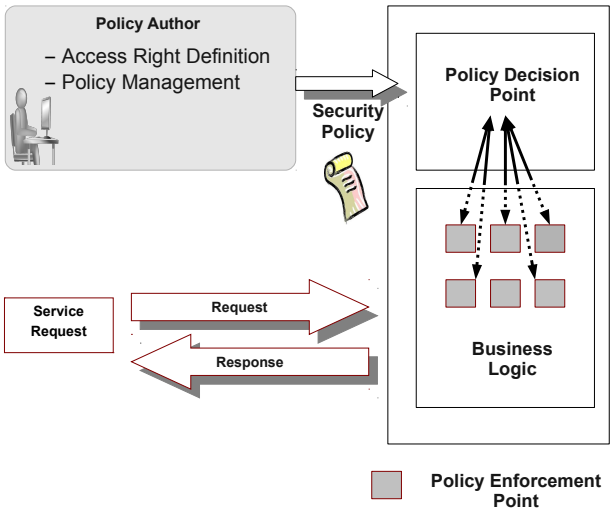


Figure 1: Access Control Request Evaluation

The request evaluation time depends on the size (number of rules) of the policy that the PDP encapsulates. For a given policy size, the evaluation time to evaluate requests increases linearly with the workload (i.e., the number of requests). Our *Hypothesis 1* is that the more rules a policy contains, the higher the slope of the evaluation time with an increasing workload. *Hypothesis 1* validity is discussed in Section 4. As a consequence, one possibility to improve performance consists in splitting the centralized PDP into PDPs with smaller policy sizes. We consider keeping the same input channel in the decentralized architecture. Therefore, we do not change the PEP code. Note that if a specific input channel is required for each PEP, developers are required to change the PEP code to map each PEP with its corresponding PDP.

2.3 Centralization: PEPs and PDP Synergy

Centralization offers a desirable feature by simplifying the routing of requests to the right PDP. Figure 2 illustrates the model of the access control architecture. In this model, a set of business processes, which comply to users' needs, is illustrated by the business logic, which is enforced by multiple PEPs. Conceptually, the decision is decoupled from the enforcement and involves a decision making process in which each PEP interacts with the same single PDP. The key point concerns the cardinality linking PEPs to the PDP. While a PDP is potentially linked to many PEPs, any PEP is strictly linked to exactly one PDP (which is unique in the centralized model). Since there is only one PDP, the requests are all routed to this unique PDP. No particular treatment is required to map a given PEP in the business logic to the corresponding PDP, embedding the requested rules. Another advantage of this many-to-one association is the clear traceability between what has been specified by the policy and the PEPs enforcing this policy at the business logic level. In such setting, when access control policies are updated or removed, the related PEPs can be easily located and updated or removed. Thus the application is updated synchronously with the policy changes. We call this desirable property *synergy* of the access control architecture: an access control architecture is said to be *synergic* if any PEP always sends its requests to the same PDP. As a consequence, splitting the centralized PDP into PDPs of smaller policy sizes may break this synergy since calls issued by PEPs can be handled by several PDPs. In this work, we consider various splitting criteria to transform a centralized PDP into PDPs with smaller policy size. Our *Hypothesis 2* is with comparable PDP p sizes, the evaluation time will be reduced when the architecture is synergic. This hypothesis is investigated in Section 4.

2.4 Tradeoff for Refactoring

The following facts are taken into account in our work:

- Access control architectures are centralized with a unique PDP.
- Centralization eases reconfiguration of an access control policy.
- Centralization threatens performance.
- Direct mapping from any PEP to only one PDP makes the access control architectures synergic.
- A synergic system facilitates PEP request routing and eases policy maintenance.

The goal of our work is to improve performance by refactoring the centralized model into its corresponding decentralized model with multiple PDPs. The resulting architecture must have an equivalent behavior and should not impact the desirable properties of the centralized model, namely reconfigurability and synergy. Automating the transformation from a centralized to a decentralized architecture is required to preserve reconfigurability. With automation, we can still reconfigure the centralized policy, and then automatically refactor the architecture. We propose automatic refactoring of a centralized model into its corresponding decentralized model while preserving high reconfigurability. However, refactoring the architecture by splitting the centralized PDP into smaller ones may break the initial synergy. This phenomenon is studied in the empirical study of Section 4 together with *Hypothesis 2*. In the next section, we give an overview of the XACML language since it is the standard language used in this paper to implement a PDP.

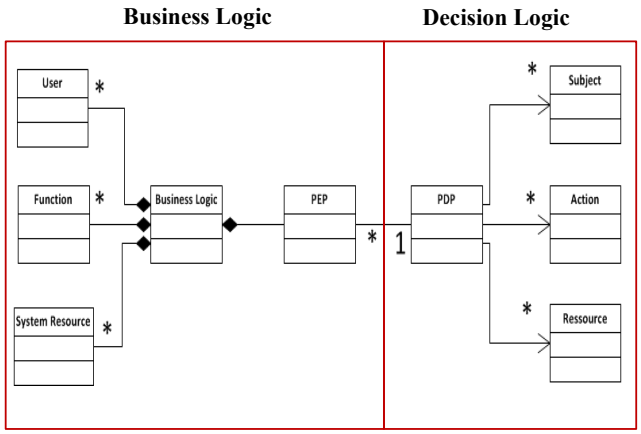


Figure 2: Access Control Model

2.5 XACML Policies and Performance Issues

In this paper, we focus on access control policies specified in the eXtensible Access Control Modeling Language (XACML) [3]. XACML is an XML-based standard policy specification language that defines a syntax of access control policies and requests/responses. XACML enables policy authors to externalize access control policies for the sake of interoperability since access control policies can be designed independently from the underlying programming language or platform. Such flexibility enables to easily update access control policies to comply with new requirements. An XACML policy is constructed as follows. A `policy set` element consists of a sequence of `policy` elements, a combining algorithm, and a `policy target` element. A `policy` element is expressed through a `target`, a set of `rules`, and a rule combining algorithm. A `target` element consists of the set of subjects, resources, and actions to which a policy or a rule is applicable. A `rule` consists of a `target` element, a `condition` element, and an `effect`. A `condition` element is a boolean expression that specifies the environmental context (e.g., time and location restrictions) in which the rule applies. Finally, an `effect` is the rule's authorization decision, which is either permit or deny. Given a request, a PDP evaluates the request against the `rules` in the policy by matching subjects, resources, and actions in the request. More specifically, an XACML request encapsulates attributes that define which subject requests to take action on which resource (e.g., subject Bob requests to borrow a book). Given a request that satisfies the `target` and `condition` elements in a rule, the rule's effect is taken as the decision. If the request does not satisfy the `target` and `condition` elements in any rule, its response yields the "NotApplicable" decision.

When more than one rule is applicable to a request, the combining algorithm helps determine which rule's effect can be finally given as the decision for the request. For example, given two rules that are applicable to the same request and provide different decisions, the permit-overrides algorithm prioritizes a permit decision over the other decisions. More precisely, when using the permit-overrides algorithm, the policy evaluation produces one of the following three decisions for a request:

- Permit if at least one permit rule is applicable for the request.
- Deny if no permit rule is applicable and at least one deny rule is applicable for the request.
- NotApplicable if no rule is applicable for the request.

A `policy target` element describes what the policy applies to by referring to attributes of subjects, resources, and actions. Figure 3 shows a simplified XACML policy that denies subject Bob to borrow a book.

XACML policies become more complex when handling increasing complexity of organizations in terms of structure, relationships, activities, and access control requirements. In such a situation, a policy often consists of a large number of rules to specify policy behaviors for various resources, users, and actions in the organizations. In policy-based systems, policy authors manage a centralized and a single PDP loaded with a single policy to govern all system resources. However, due to a large number of rules for evaluation, this centralization raises performance concerns related to request evaluation time for access control policies and may degrade the system efficiency and slow down the overall business processes.

We present the following three main factors that may cause to degrade XACML request evaluation performance:

- An XACML policy may contain various attribute elements including `target` elements. Retrieval of attribute values in the `target` elements for request evaluation may increase the evaluation time.
- A `policy set` consists of a set of policies. Given a request, a PDP determines the final authorization decision (i.e., effect) of the whole `policy set` after combining all the applicable rules' decisions for the request. Computing and combining applicable rules' decisions contribute to increasing the evaluation time.
- `Condition` elements in rules can be complex because these elements are built from an arbitrary nesting of boolean functions and attributes. In such a situation, evaluating `condition` elements may slow down request evaluation time.

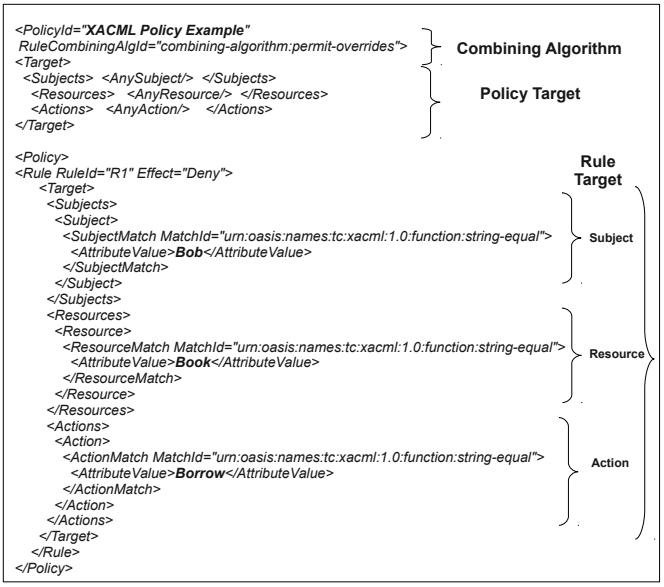


Figure 3: XACML Policy Example

3. POLICY REFACTORING

This section describes our approach of refactoring access control policies to improve performance by reducing the number of policy rules potentially applicable to a request. For refactoring policies in a systematic way, we propose seven policy splitting criteria based on attribute sets. Moreover, we explain how to select a splitting criterion that preserves the synergy in the access control architecture.

3.1 Policy Splitting Criteria

During the evaluation process, the attribute values in a given request are compared with the attribute values in the target of a rule. If there is a match between the request's attribute values and target's attribute values, the rule is then applicable to the request. In the decision making process, applicable rules contribute to determining the final authorization decision whereas non-applicable rules are not relevant in this process. For request evaluation, not all the rules are applicable to the request. In other words, only part of the rules (i.e, relevant rules) are applicable to the request and can contribute to determining the final decision.

We propose an approach to evaluate a request against only the relevant rules for the given request by refactoring the access control policies. Our approach aims at splitting a single global policy

into multiple smaller policies based on attribute combination. For a given policy-based system, we transform its policy P into policies P_{SC_w} containing a smaller number of rules and conforming to a Splitting Criterion SC_w. An SC_w defines the set of attributes that are considered to classify all the rules into subsets each with the same attribute values and w denotes the number of attributes that have to be considered conjointly for aggregating rules based on specific attribute elements. Table 1 shows our proposed splitting criteria categorized according to attribute element combinations.

Table 1: Splitting Criteria

Categories	Splitting Criteria
SC_1	$\langle Subject\rangle, \langle Resource\rangle, \langle Action\rangle$
SC_2	$\langle Subject, Action\rangle, \langle Subject, Resource\rangle$ $\langle Resource, Action\rangle$
SC_3	$\langle Subject, Resource, Action\rangle$

To illustrate our approach, we present examples that take into consideration the XACML language features. In Figure 4, our approach refactors an XACML policy P according to the splitting criterion $SC_1 = \langle Subject\rangle$. Our refactoring results in two sub-policies Pa and Pb. Each sub-policy consists of relevant rules with regards to the same subject (Alice or Bob in this case).

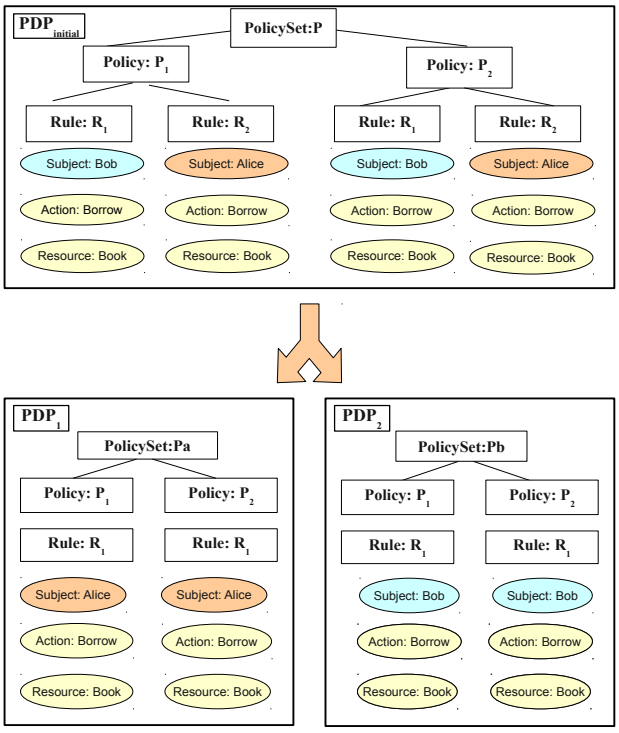


Figure 4: Refactoring a Policy According to $SC_1 = \langle Subject\rangle$

Technically, to split a given policy P according to $SC_1 = \langle Subject\rangle$, we start by parsing the global policy P and by collecting the overall subject attribute values in the policy. For each collected subject attribute value Sa, we consider the global policy and we delete the rules that do not contain Sa as a subject attribute value in the target element attributes. After all the successive deletions, the global policy is refactored to a policy that contains only the rules with Sa in their subject attribute values. Algorithm 1 describes the splitting process for $SC_1 = \langle Subject\rangle$.

Algorithm 1 Policy Splitting Algorithm for $SC_1 = \langle Subject\rangle$

Input: XACML Policy P, Splitting Criterion SC_1=$\langle Subject\rangle$
Output: Sub-policies Set: S
SplitPolicy()
S=Ø
/* Collect all subjects in all the rules /*
for each Rule R_i in Policy P **do**
  /* Fetch all the targets to extract attribute collection based on SC */
  for each Target.Subject in R_i **do**
    SubjectCollection.add(SubjectElement.attribute)
  end for
end for
/* Build sub-policies based on subjects collected in SubjectCollection */
for int $i = 0$; $i <$ SubjectCollection.size(); $i++$ **do**
  /* Remove all the rules that do not contain SubjectCollection.at(i) in their Target */
  for each Rule R_i in Policy P **do**
    if $R_i.Target.SubjectElement$! = AnySubject **then**
      if (Target.SubjectElement.attribute in R_i) ! = SubjectCollection.at(i) **then**
        Remove R_i
      end if
    end if
  end for
  /* $P_{(SubjectCollection.at(i))}$ is a sub-policy with only rules where the subjectAttribute is equal to SubjectCollection.at(i) */
  /* Add the sub-policy to the set of sub-policies */
  $S = S \cup P_{(SubjectCollection.at(i))}$
end for

Our algorithm is safe in the sense that it does not change the authorization behavior of the PDP. There are two important issues to be considered when reasoning about the safety of the algorithm:

- Can the splitting impact authorization results when a policy set includes multiple policies with different combining algorithms?
- When AnySubject, AnyAction or AnyResource is used as target element values, does the splitting change the behavior of the PDP?

The first issue is addressed by the way the algorithm operates. The first step of the algorithm goes through all the rules and extracts the set of target element values (the set of subjects, the set of actions, and/or the set of resources) based on the splitting criterion. Then, based on the extracted result, the splitting is performed by removing the rules with different splitting criterion values (such as a subject different from the splitting criterion subject). The rules that are kept are therefore not modified and their behavior is not altered. When there are several policies with different combining algorithms, the rules that are kept do not impact the evaluation behavior because they remain attached to the same combining algorithm. Moreover their order and their content are not modified.

The second issue is addressed by keeping all the rules that involve AnySubject, AnyAction, or AnyResource in all sub-policies

because by definition during evaluation, these values are taken into consideration for evaluating all possible values of subjects, actions, and resources. It is worth mentioning this following consideration related to the refactoring process: XACML supports multi-valued attributes in policies and requests. In XACML policies, `target` elements define a set of attribute values, which match with the context element in an access control request. In Figure 5, the subject attribute includes two attributes (one is "role" and the other is "isEq-subjUserId-resUserId"). In order to match the subject with multi-valued attributes, a request should include at least `pc-member` and `true` for "role" and "isEq-subjUserId-resUserId", respectively. Our approach considers such a whole subject element as a single entity, which is not split by the policy splitter component.

```
<Subjects>
  <Subject>
    <SubjectMatch MatchId="urn:oasis:names:tc:xacml:1.0:function:string-equal">
          <AttributeValue>Administrator</AttributeValue>
    <SubjectAttributeDesignator AttributeId="role"/>
    </SubjectMatch>
    <SubjectMatch MatchId="urn:oasis:names:tc:xacml:1.0:function:string-equal">
          <AttributeValue >true</AttributeValue>
    <SubjectAttributeDesignator AttributeId="isEq-subjUserId-resUserId"/>
  </SubjectMatch>
  </Subject>
</Subjects
```

Figure 5: Multi-attribute Values in `target` Element

After the splitting is performed, our approach creates one or more PDPs that comply with the splitting criterion. We use the Sun PDP [3] to evaluate a request against policies specified in XACML. During request evaluation, the Sun PDP checks the request against the policy and determines whether the decision is permit or deny. Given a request, our approach runs the Sun PDP loaded with the request's relevant policy, which is used during the decision making process. The PDP then retrieves the rules that are applicable to the request. Figure 6 presents our approach to handling request evaluation with multiple policies. During the evaluation process, given a request, our approach verifies the matching between the request's attribute value and the policy target elements attributes. Our approach then selects only the relevant policy among all the policies for a given request. After the selection of the relevant policy, all of its relevant rules for the decision making are evaluated. Figure 7 shows an overview of our approach. In our approach, the policy splitter component plays a role to refactor access control policies. Given a single PDP loaded with the initial global policy, the policy splitter component conducts automated refactoring by creating multiple PDPs loaded with XACML policies, which are split from the initial global policy based on the user-specified splitting criterion. If the initial global policy is changed, the policy splitter component is required to refactor the policy again to create PDPs with the most recent relevant policies. Our refactoring approach is safe in the sense that the approach does not impact existing security aspects in a given system.

3.2 Architecture Model Preservation: PEP-PDP Synergy

We propose to preserve the synergy property in the access control architecture by mapping a PEP and a PDP loaded with the relevant policy for a request dynamically at runtime. As shown in Section 3.1, given multiple PDPs after the policy refactoring, we consider (1) how PEPs are organized at the application level, and (2) how PEPs are linked to their corresponding PDPs. In the worst case, splitting the initial PDP into multiple PDPs may lead to a non-

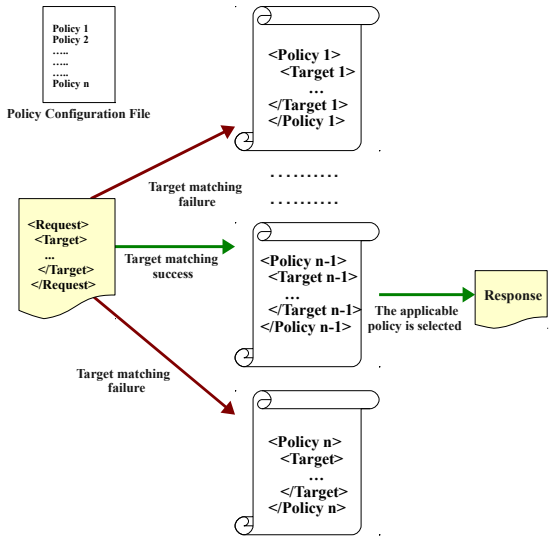


Figure 6: Applicable-Policy Selection

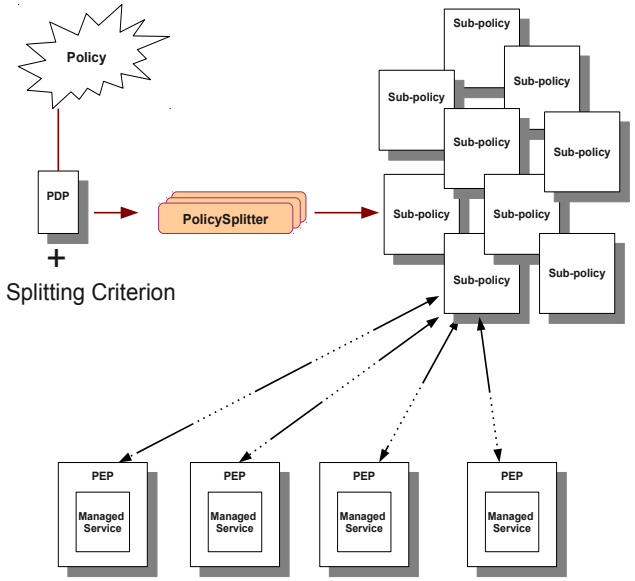


Figure 7: Overview of the Refactoring Process

synergic system: a PEP may send its requests to several PDPs. The PDP that handles a given request is only known at runtime. Such a resulting architecture breaks the PEP-PDP synergy and the conceptual simplicity of the initial architecture model. In the best case, the refactoring preserves the simplicity of the initial architecture by keeping a many-to-one association between PEPs to PDPs. Given a request, our approach maps a PEP to a PDP with relevant rules for the request. Therefore, different requests issued from a PEP should be handled by the same PDP. Operationally, the request evaluation involves one policy. In this case, our refactoring does not impact the conceptual architecture of the system.

Figure 8 presents a PDP encapsulating a global policy that has

been refactored. The system that is presented on the left is resulted from a desirable refactoring whereas the one on the right is resulted from an undesirable refactoring.

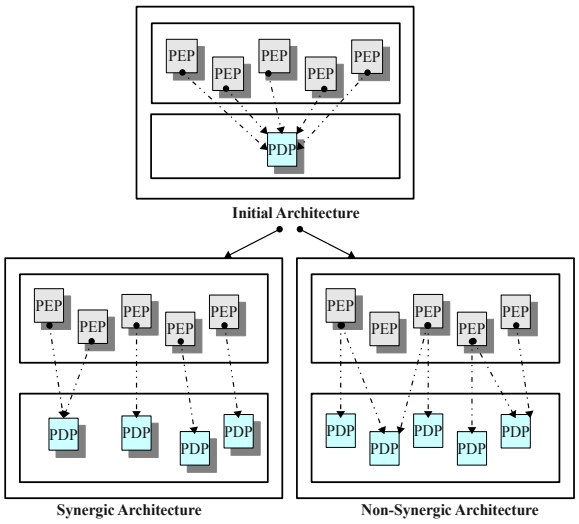


Figure 8: Synergic vs Non-synergic System

At the application level, a PEP is represented by a method call that triggers a decision making process. Figure 9 presents a sample PEP code snippet from our previous work [11]. This code snippet shows an example of a PEP represented by the method `checkSecurity`, which calls a method of the class `SecurityPolicyService`, which formulates a request to invoke the PDP component. The PEP represented by the method `ServiceUtils.checkSecurity` may issue requests that have subject "user" along fixed action and resource ("LibrarySecurityModel.BORROWBOOK_METHOD"), ("LibrarySecurityModel.BOOK_VIEW"). Consider that we refactor a policy using $SC_2 = \langle Resource, Action \rangle$, $SC_1 = \langle Action \rangle$, or $SC_1 = \langle Resource \rangle$. Given a request issued from the PEP, our approach runs a PDP loaded with a policy containing rules sharing the same action and resource attribute values. Thus the splitting process that preserves the mapping between the PEPs and the PDP is the one that considers the following splitting criteria: $SC_2 = \langle Resource, Action \rangle$, $SC_1 = \langle Action \rangle$, and $SC_1 = \langle Resource \rangle$ in this case. In the evaluation section, we investigate the impact of the synergy property on performance.

```
public void borrowBook(User user, Book book) throws
SecuritPolicyViolationException {

// call to the security service
        ServiceUtils.checkSecurity(user,
LibrarySecurityModel.BORROWBOOK\_METHOD,
LibrarySecurityModel.BOOK\_VIEW),
ContextManager.getTemporalContext());}

  // call to business objects
  // borrow the book for the user
book.execute(Book.BORROW, user);
  // call the dao class to update the database
bookDAO.insertBorrow(userDTO, bookDTO);}
```

Figure 9: PEP Deployment Example

4. EVALUATION

We carried out our evaluation on a desktop PC running Ubuntu 10.04 with a Core i5, 2530 Mhz processor, and 4 GB of RAM. We have implemented a tool, called `PolicySplitter` to split policies according to a given splitting criterion automatically. The tool is implemented in Java and is available for download [4].

4.1 Objectives and Metrics

Our evaluation intends to answer the following research questions:

1. **RQ1.** How faster can request evaluation time of multiple Sun PDPs with policies split by our approach achieve compared to that of an existing single Sun PDP? This question helps show that our approach can improve performance in terms of request evaluation time. Moreover, we compare request evaluation time for different splitting criteria.

2. **RQ2.** With comparable PDP policy sizes, is request evaluation time of a system faster when its architecture is synergic? This research question investigates *Hypothesis 2* presented in Section 2.

3. **RQ3.** How faster can request evaluation time of XEngine with policies split by our approach achieve compared to that of an existing single XEngine? This question helps show that our approach can improve performance in terms of request evaluation time for other advanced policy evaluation engines such as XEngine.

4. **RQ4.** How faster does request processing time of XEngine with policies split using our technique achieve compared to that of the Sun PDP with policies split using our technique? This question aims at checking whether XEngine in combination with our technique performs better than the Sun PDP combined with our technique as well.

5. **RQ5.** For larger PDP policy size, do we observe higher slope of the evaluation time with an increasing workload? This research question investigates *Hypothesis 1* (presented in Section 2) on the impact of the number of rules in a given PDP on the evaluation time.

To address these research questions, we go through the following evaluation setup based on two different empirical studies:

- First, we evaluate the performance improvement regarding the decision making process by taking into consideration the whole system (PEPs and PDPs). We compared request evaluation time with a single global policy (handled by a single PDP) against request evaluation time with split policies. All the splitting criteria have been considered in our evaluation. `IA` denotes an "Initial Architecture", which uses the single global policy for request evaluation. This step allows studying the behavior of splitting criteria that preserve the synergy property in the access control architecture.

- Second, we evaluate the performance of PDPs in isolation to compare the splitting criteria independently from the system. Recall that in this step, we do not reason about the synergy property since we do not consider the application level. We replace Sun PDPs with XEngine [13] to investigate the effectiveness of our approach with other decision engines.

4.2 Subjects

The subjects include three real-life Java systems each of which interacts with access control policies. Full details on our subjects are available [11, 17, 18]. We next describe our three subjects.

- The Library Management System (LMS) provides web services to manage books in a public library.
- The Virtual Meeting System (VMS) provides web conference services. VMS allows users to organize online meetings in a distributed platform.
- The Auction Sale Management System (ASMS) allows users to buy or sell items online. A seller initiates an auction by submitting a description of an item that she wants to sell with its expected minimum price. Users then participate in the bidding process by bidding the item. To bid on the item, user must have enough money in his/her account before bidding.

Our subjects are initially built upon the Sun PDP [3] as a decision engine, which is a popularly used PDP to evaluate requests. We started by a processing step, in which we have augmented the rules in the three original policies for these studies, as it would be difficult to observe performance improvement results with systems including few rules. In our evaluations, LMS policy contains 720 rules, VMS has 945 rules, and ASMS has 1760 rules. The rules that we added do not modify the system behavior as they are conform to the specifications. Moreover, to assess performance improvement over an existing advanced PDP, we adopt XEngine (instead of the Sun PDP) in our subjects to evaluate requests. XEngine is an advanced policy evaluation engine, which transforms the hierarchical tree structure of the XACML policy to a flat structure to reduce request evaluation time. XEngine also handles various combining algorithms supported by XACML.

4.3 Performance Improvement: Sun PDP

In order to answer **RQ1**, we generated the resulting sub-policies for all the splitting criteria defined in Section 3.1. For each splitting criterion, we have executed system tests to generate requests that trigger all the PEPs in the evaluation. The test generation step leads to the execution of all combinations of possible requests described in our previous work [18]. The process of test generation is repeated ten times to alleviate the impact of randomness. We applied this process for each splitting criterion and calculated evaluation time on average of a system under test. Figure 10 presents evaluation time for policies split based on each splitting criterion and the global policy of the subjects. We can make two observations:

- Compared to the evaluation time of `IA`, our approach improves performance for all of splitting criteria in terms of evaluation time. This observation is consistent with our expected results; the evaluation time against policies with a smaller number of rules (compared with the number of rules in `IA`) is faster than that against policies in `IA`.
- The splitting criterion $SC = \langle Action, Resource \rangle$ enables to show the fastest evaluation time. Such observation is due to the fact that the PEPs in our three subjects are organized based on $SC_2 = \langle Resource, Action \rangle$. This observation pleads in favor of applying a splitting criterion that takes into account the PEP-PDP synergy.

To identify the splitting criterion that generates the smallest number of PDPs, we have studied the number of policies generated by the splitting. Figure 11 shows the results. We observed the number of policies based on our proposed three categories: (1) the SC_1 category leads to the smallest number N_1 of PDPs, (2) the SC_2 category leads to a medium number N_2 (N_1<N_2<N_3) of PDPs, and (3) SC_3 leads to the largest number N_3 of PDPs. While the SC_1 category leads to the smallest number of PDPs, each PDP encapsulates a relatively large number of rules in a policy (compared with that of SC_2 and SC_3, which leads to performance degradation). We have classified splitting criteria according to their preservation of the synergy property considering our subjects. The classification is shown in Table 2 where S denotes Subject, R denotes Resource, A denotes Action, and IA denotes Initial Architecture. For example, AR denotes SC=$< Action, Resource >$. AR, A, and R are synergic splitting criteria since all the PEPs in our considered three systems are organized as shown in Figure 9.

Table 2: Splitting Criterion Classification

	S	A	R	SA	SR	AR	SAR	IA
Synergic		x	x			x		x
Not-Synergic	x			x	x		x	

To answer **RQ2**, we have evaluated PDPs in the three systems and for the different splitting criteria. The results presented in Figure 12 show the average number of rules in each PDP, for each splitting criterion in the three systems. We can observe that the AR criterion produces comparable size of PDPs with the SR criterion; however, as shown in Figure 10, AR is the best splitting criterion in terms of evaluation time performance. Moreover, the number of PDPs produced with the splitting critera S and A is comparable; the criterion A, which is synergic, has evaluation time less than the one produced by the splitting criterion S, which is not synergic. This result supports our *Hypothesis 2*, which states that with comparable PDP sizes, the evaluation time would be reduced when the architecture is synergic.

4.4 Performance Improvement: XEngine

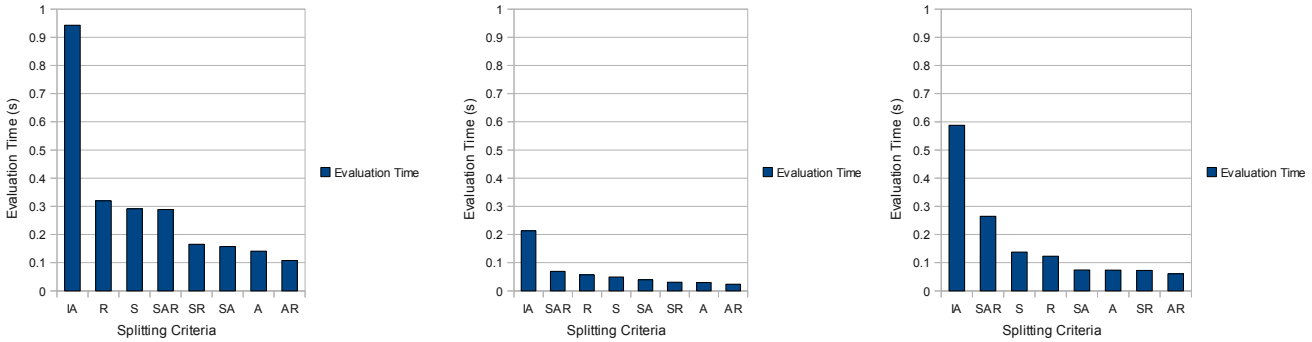
In order to answer **RQ3** and **RQ4**, we measure request evaluation time of XEngine with policies split by our approach compared with that of an existing single XEngine. The goal of this empirical study is to show the impact of combining XEngine with our splitting process. XEngine itself improves dramatically the performance of the SUN PDP mainly for three reasons:

- It uses a refactoring process that transforms the hierarchical structure of the XACML policy to a flat structure.
- It converts multiple combining algorithms to single one.
- It relies on a tree structure that minimizes the request evaluation time.

We propose to use XEngine conjointly with the refactoring process presented in this work. We have evaluated our approach in two settings:

- Considering evaluation with a decision engine based on XEngine with split policies and with the initial policy.
- Considering evaluation with a decision engine based on Sun PDP with split policies and with the initial policy.

In this step, we do not reason about the synergy, since we do not consider the application level for the three systems. We measure request evaluation time by evaluating a randomly generated set of 10,000 requests as proposed in our previous work [15]. The request



(a) LMS (b) VMS (c) ASMS

Figure 10: Request Evaluation Time for the Three Subjects

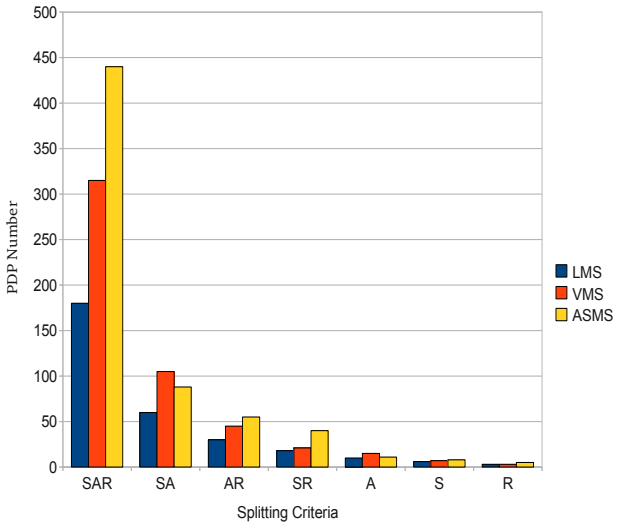


Figure 11: PDP Number Produced with Splitting Criteria

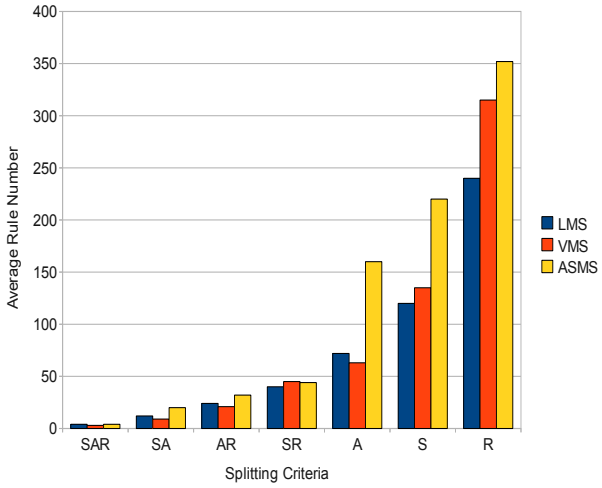


Figure 12: Average of Rule Numbers per PDP in the Three Systems

evaluation time is evaluated for the three systems. The results are presented in Tables 3, 4, and 5. In the three tables, the percentage of performance improvement % PI shows the gain in request evaluation time with the policies split by our approach compared the initial architecture (IA) with SUN PDP and XEngine, respectively.

XEngine with split policies, in most cases, enables to reduce the evaluation time compared to XEngine with a single policy. This result is shown in Table 5 for ASMS where the evaluation time is reduced about 33 times from 1639 ms in the initial architecture (IA) to 49 ms with the splittig criterion SAR. This empirical observation shows that our refactoring conjointly with XEngine enables to improve the performance of the evaluation process for most of the splitting criteria and thus answers **RQ3**.

As shown in the three tables, there are some splitting criteria which lead to decrease of performance like the splitting criterion R in VMS system which leads to decrease the evlaution time to -44%. These results need to be investigated with other case studies.

Through the three tables, we observe that, when our subjects are equipped with XEngine, our proposed approach substantially improves performance (compared to the results with the Sun PDP) for most of the splitting criteria. For the splitting criteria SC=$\langle Action \rangle$ abbreviated as A, in the LMS system, the evaluation time is reduced about 22 times: from 2703 ms to 120 ms with XEngine, this observation enables to answer **RQ4**.

4.5 Impact of Increasing Workload

To investigate **RQ5**, we have calculated request evaluation time according to the number of requests incoming to a system. For each policy in the three systems (ASMS, LMS, and VMS), we generated 5000, 10000, .., 50000 random requests to measure the evaluation time (ms). The results are shown in Figure 13. For the three systems, we observe that the evaluation time increases when the number of requests increases in a system. With an increasing system load, the request evaluation time is considerably improved when using the splitting process compared to the initial architecture. The results shown in Figure 13 are interpreted by the average of PDP sizes presented in Figure 12. The results are consistent with *Hy-*

Table 3: Evaluation Time (ms) in LMS

	SAR	AR	SA	SR	R	S	A	IA
Sun PDP	485	922	1453	1875	2578	2703	2703	2625
% PI	81.5	64.9	44.6	28.6	1.2	-3	-3	0
XEngine	26	47	67	95	190	164	120	613
% PI XEngine	97.7	92.3	89.0	84.5	69	73.2	80.4	0

Table 4: Evaluation Time (ms) in VMS

	SAR	AR	SA	SR	R	S	A	IA
Sun PDP	1281	2640	3422	3734	6078	5921	6781	5766
% PI Sun PDP	77.8	54.2	40.6	35.2	5.4	-2.7	-17.6	0
XEngine	34	67	96	145	384	274	149	265
% PI XEngine	87.2	74.7	63.8	45.3	-44.9	-3.4	43.8	0

Table 5: Evaluation Time (ms) in ASMS

	SAR	AR	SA	SR	R	S	A	IA
Sun PDP	2280	2734	3625	8297	7750	8188	6859	7156
% PI Sun PDP	68.1	61.8	49.3	15.9	-8.3	-14.4	4.1	0
XEngine	49	60	104	196	310	566	262	1639
% PI XEngine	97	96.3	93.65	88	81	65.5	84	0

pothesis 1 (presented in Section 2), which states that the slope of evaluation time increases with PDP size in a system with an increasing workload.

To deploy our approach, we need to fetch the relevant PDP for a given request at runtime. Therefore, request processing time includes both fetching time and request evaluation time. Figure 14 shows percentage of fetching time over the global evaluation time for request evaluation in LMS. The fetching time increases according to the PDP size. The fetching time is relatively small in comparison with the total evaluation time and thus does not impact significantly the evaluation time.

4.6 Summary

We summarize the results of the evaluation:

- We have experimentally shown the effectiveness of the splitting in reducing the evaluation time. Our refactoring process improves both a typical PDP (the SUN PDP) and an advanced PDP (XEngine).
- When the sizes of PDPs are comparable, the splitting criteria that are synergic enable to have the best results in terms of evaluation time.

The evaluation of the synergy property on improving performance has to be strengthened by conductinng other experiments on other evaluation subjects and by considering different organizations of PEPs at the application level.

4.7 Threats to Validity

The threats to external validity primarily include the degree to which subjects, policies, and test requests are representative of true practice. These threats could be reduced by further evaluation on a wider type and larger number of policies and a larger number of test requests in future work. In particular, our approach is based on only seven proposed splitting criteria. We could develop additional splitting criteria to split policies and measure efficiency in terms of request evaluation time. In addition, our approach generates random test requests, which may induce bias or randomness in our results. To prevent such a bias, we conduct our evaluation for 10 times and measure an average value of evaluation results. The threats to internal validity are instrumentation effects that can bias our results such as faults in the Sun PDP, XEngine, `PolicySplitter`, measurement tool in terms of request evaluation, and random request generators.

5. RELATED WORK

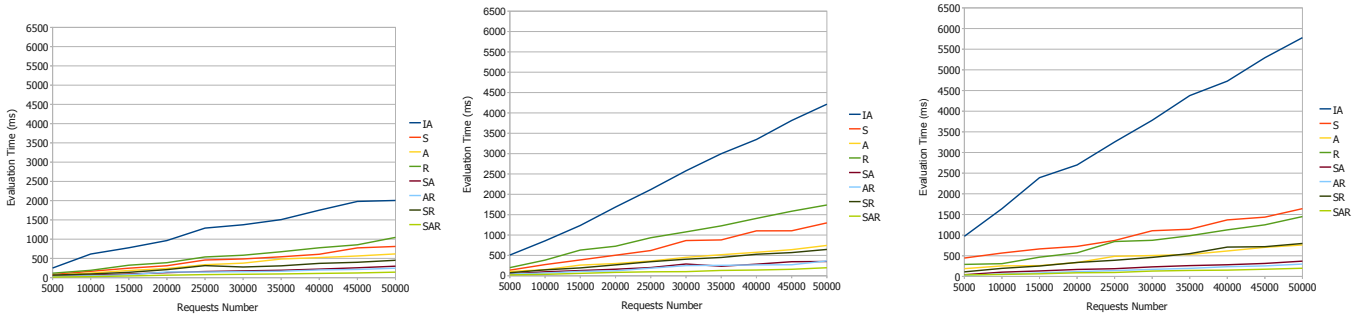
There are several previous approaches about performance issues in security mechanisms. Ammons et al. [5] have presented techniques to reduce the overhead engendered from implementing a security model in IBM's WebSphere Application Server (WAS). Their approach identifies bottlencks through code instrumentation and focuses on two aspects: the temporal redundancy (when security checks are made frequently) and the spatial redundancy (using the same security techniques on the same code execution paths). For the first aspect, they use caching mechanisms to store checks results, so that the decision is retrieved from the cache. For the second aspect, they used a technique based on specialization, which consists in replacing an expensive check with a cheaper one for frequent paths. While this previous approach focuses on bottlencks in program code, in this paper, we propose a new approach to refactor access control policies by reducing the number of rules in each split policy.

Various approaches [8,12,14] have been proposed to address performance issues in systems interacting with access control policies. Jahid et al. [8] focus on XACML policy verification for database access control. They presented a model that converts attribute-based policies into access control lists. They implemented their approach called MyABDAC. While they measured performance of MyABDAC in terms of request evaluation, they did not show how much MyABDAC gains improvement over an existing PDP.

Marouf et al. [14] have proposed an approach for policy evaluation based on a clustering algorithm that reorders rules and policies within the policy set so that the access to applicable policies is faster. Their categorization is based on the subject target element. Their approach requires identifying the rules that are frequently used. Our approach follows a different strategy and does not require knowing which rules are used frequently. In addition, the rule reordering is tightly related to specific systems. If the PDP is shared between several systems, their approach could not be applicable since the most "used" rules may vary between systems.

Lin et al. [12] decomposed a global XACML policy into local policies related to collaborating parties, and the local policies are sent to corresponding PDPs. The request evaluation is based on local policies by considering the relationships among local policies. In their approach, the optimization is based on storing the effect of each rule and each local policy for a given request. Caching decision results is then used to optimize evaluation time for an incoming request. However, there were no experimental results for measuring the efficiency of their approach when compared to the traditional architecture. While the previous approaches have focused on the PDP component to optimize the request evaluation, Miseldine et al. [16] addressed this problem by analyzing rule location on XACML policies and requests at the design level so that the relevant rules for the request are accessed faster on evaluation time.

Our contribution in this paper brings new dimensions over our previous work on access control [13, 17, 18]. We have proposed XEngine [13], which focuses particularly on performance issues addressed with XACML policy evaluation. XEngine proposes an alternative solution to brute force searching based on an XACML policy conversion to a tree structure to minimize the request eval-



(a) LMS (b) VMS (c) ASMS

Figure 13: Evaluation Time for Our Subjects, LMS, VMS, and ASMS Depending on the Request Number

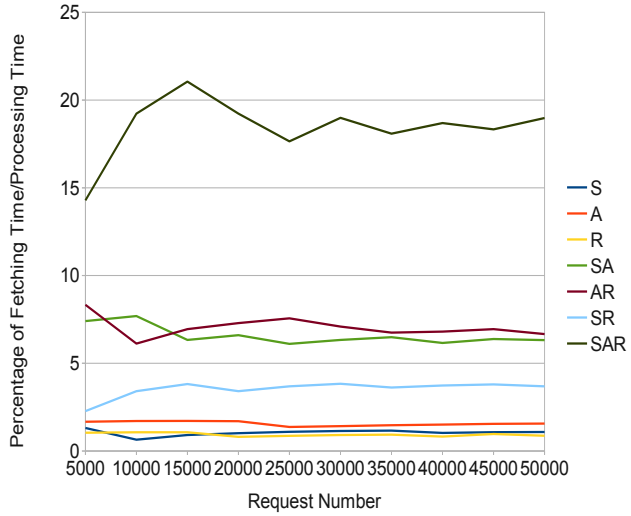


Figure 14: Percentage of Fetching Time

uation time. It involves a refactoring process that transforms the global policy to a decision diagram that is then converted to forwarding tables. In our contribution in this paper, we introduce a new refactoring process that involves splitting the policy into smaller sub-policies. Our two refactoring processes are combined to decrease dramatically the evaluation time.

6. CONCLUSION AND FUTURE WORK

In this paper, we have tackled the performance issue in the decision making mechanism for access control and have proposed an automated refactoring process that enables to reduce request evaluation time substantially. All the reasonings about performance factors have not included hardware considerations. However, performance improvement at the software/logical level as done by our approach would complement performance improvement at the hardware level to best improve the overall performance. Our approach has been applied to XACML policies and it can be generalized to policies in other policy specification languages (such as EPAL). To support and automate the refactoring process, we have designed and implemented the PolicySplitter tool, which transforms a given policy into small ones, according to a chosen splitting criterion. Most obtained results have shown a significant gain in evaluation time. The best gain in performance is reached by the criterion that respects the synergy property. This result pleads in favor of a refactoring process that takes into account the way PEPs are scattered inside the system business logic. In this work, we have easily identified the different PEPs since we know exactly how our system functionalities are implemented and thus how PEPs are organized inside the system. In future work, we plan to automatically identify the different PEPs of a given system. This technique is an important step complementary to our current approach, since this technique enables knowing how PEPs are organized in the system and thus allows to select automatically the most suitable splitting criterion for a given system.

7. ACKNOWLEDGMENTS

This work is supported in part by NSF grants CCF-0845272, CCF-0915400, CNS-0958235, an NIST grant, and ARO Grant No. W911NF-08-1-0443.

8. REFERENCES

[1] IBM, Enterprise Privacy Authorization Language (EPAL), Version 1.2 . http://www.w3.org/Submission/2003/SUBM-EPAL-20031110, 2003.

[2] OASIS eXtensible Access Control Markup Language (XACML). http://www.oasis-open.org/committees/xacml/, 2005.

[3] Sun's XACML implementation. http://sunxacml.sourceforge.net/, 2005.

[4] PolicySplitter Tool. http://www.mouelhi.com/policysplitter.html, 2011.

[5] G. Ammons, J. deok Choi, M. Gupta, and N. Swamy. Finding and removing performance bottlenecks in large systems. In *Proceedings of European Conference on Object-Oriented Programming, ECOOP*. Springer, 2004.

[6] E. D. Bell and J. L. La Padula. Secure computer system: Unified exposition and multics interpretation. Mitre Corporation, 1976.

[7] D. F. Ferraiolo, R. S. Sandhu, S. I. Gavrila, D. R. Kuhn, and R. Chandramouli. Proposed NIST standard for role-based access control. *ACM Transactions on Information and System Security*, 4(3):224–274, 2001.

[8] S. Jahid, C. A. Gunter, I. Hoque, and H. Okhravi. Myabdac: compiling xacml policies for attribute-based database access control. In *Proceedings of the first ACM conference on Data and application security and privacy*, pages 97–108, 2011.

[9] A. A. E. Kalam, S. Benferhat, A. Miège, R. E. Baida, F. Cuppens, C. Saurel, P. Balbiani, Y. Deswarte, and G. Trouessin. Organization based access contro. In *Proceedings of 10th IEEE International Conference on Policies for Distributed Systems and Networks*, pages 120–131, 2003.

[10] B. Lampson. Protection. In *Proceedings of the 5th Princeton Conference on Information Sciences and Systems*, 1971.

[11] Y. Le Traon, T. Mouelhi, A. Pretschner, and B. Baudry. Test-driven assessment of access control in legacy applications. In *Proceedings of the 2008 International Conference on Software Testing, Verification, and Validation*, pages 238–247, 2008.

[12] D. Lin, P. Rao, E. Bertino, N. Li, and J. Lobo. Policy decomposition for collaborative access control. In *Proceedings of the 13th ACM Symposium on Access Control Models and Technologies*, pages 103–112, 2008.

[13] A. X. Liu, F. Chen, J. Hwang, and T. Xie. Xengine: A fast and scalable XACML policy evaluation engine. In *Proceedings of International Conference on Measurement and Modeling of Computer Systems*, pages 265–276, 2008.

[14] S. Marouf, M. Shehab, A. Squicciarini, and S. Sundareswaran. Statistics & clustering based framework for efficient xacml policy evaluation. In *Proceedings of 10th IEEE International Conference on Policies for Distributed Systems and Networks*, pages 118–125, 2009.

[15] E. Martin, T. Xie, and T. Yu. Defining and measuring policy coverage in testing access control policies. In *Proceedings of 8th International Conference on Information and Communications Security*, pages 139–158, 2006.

[16] P. L. Miseldine. Automated xacml policy reconfiguration for evaluation optimisation. In *Proceedings of 4th International Workshop on Software Engineering for Secure Systems*, pages 1–8, 2008.

[17] T. Mouelhi, F. Fleurey, B. Baudry, and Y. Traon. A model-based framework for security policy specification, deployment and testing. In *Proceedings of 11th International Conference on Model Driven Engineering Languages and Systems*, pages 537–552, 2008.

[18] T. Mouelhi, Y. L. Traon, and B. Baudry. Transforming and selecting functional test cases for security policy testing. In *Proceedings of 2009 International Conference on Software Testing Verification and Validation*, pages 171–180, 2009.

[19] R. Yavatkar, D. Pendarakis, and R. Guerin. A framework for policy-based admission control. RFC Editor, 2000.